

Sports-medical screening in middle-aged man

Citation for published version (APA):

Bovens, A. M. P. M. (1991). *Sports-medical screening in middle-aged man*. [Doctoral Thesis, Maastricht University]. Rijksuniversiteit Limburg. <https://doi.org/10.26481/dis.19911129ab>

Document status and date:

Published: 01/01/1991

DOI:

[10.26481/dis.19911129ab](https://doi.org/10.26481/dis.19911129ab)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Sports-medical screening in middle-aged man

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Bovens, Alphonsus Maximiliaan Paulus Marie

Sports-medical screening in middle-aged man / Alphonsus
Maximiliaan Paulus Marie Bovens. - Maastricht : Datawyse.
- Ill.

Thesis Maastricht. - With ref. - With summary in Dutch.

ISBN 90-5291-068-5

NUGI 468

Subject heading: sports-medical screening.

C 1991 A.M.P.M. Bovens

Omslagontwerp: Martine Bovens-Swinnen

Druk: Datawyse Maastricht / Krips Repro Meppel

Sports-medical screening in middle-aged man

proefschrift

ter verkrijging van de graad van doctor
aan de Rijksuniversiteit Limburg te Maastricht,
op gezag van de Rector Magnificus, prof. mr. M.J. Cohen,
volgens het besluit van het College van Dekanen,
in het openbaar te verdedigen
op vrijdag, 29 november 1991 om 16.00 uur

door

Alphonsus Maximiliaan Paulus Marie Bovens
geboren te Maastricht 1957

promotor: prof. dr. W.H.M. Saris

co-promotores: dr. M.A. van Baak
dr. F.T.J. Verstappen

beoordelingscommissie: prof. dr. F. Sturmans (voorzitter)
prof. dr. R.A. Binkhorst (KU-Nijmegen)
prof. dr. Tj. de Boorder
prof. dr. J.M. Greep
dr. E.F.M. Wouters.

Het verschijnen van dit proefschrift werd mede mogelijk gemaakt door de steun van het Instituut Sportgeneeskunde Limburg en het Streekge-
west Westelijke Mijnstreek.

Voor Martine
Myra
Monique

Contents

Introduction and outline of the thesis	1
<hr/>	
Chapter 1	
A medical screening in men and women over 40 years of age, active in sports	3
<hr/>	
Chapter 2	
Variability and reliability of joint measurements	17
<hr/>	
Chapter 3	
Variability and reliability of pulmonary function tests and maximal aerobic exercise testing in 3 different laboratories	29
<hr/>	
Chapter 4	
Pulmonary function in middle-aged men and women, active in sports; normal values, relation to smoking habits and sport activities	43
<hr/>	
Chapter 5	
Peak plasma lactate concentration after maximal exercise testing in people of 40 to 65 years: difference between genders	57

Chapter 6

Maximal aerobic power in cycle ergometry in middle-aged men and women, active in sports, in relation to age and physical activity	67
---	----

Chapter 7

The inter-relationship between cardiovascular fitness, physical activity and cardiovascular risk factors in middle-aged men and women, active in sports	83
---	----

Chapter 8

General discussion and conclusions	97
------------------------------------	----

Samenvatting	103
--------------	-----

Nawoord	108
---------	-----

Curriculum vitae	110
------------------	-----

Introduction and outline of the thesis

This dissertation is a reflection of part of the activities performed by the author and colleagues in the extensive population study: "The effectiveness of a sports-medical check-up in men and women over 40 years". In 1983 the preparations were started for this epidemiological study, in which the Institute of Sports Medicine Limburg collaborated with the University of Limburg and three Centers for Public Health. In the period between 1984 and 1988 almost 4000 men and women, active in sports and older than 40 years were examined during an extensive sports-medical screening. The examinations were performed by three physicians at the Centers for Public Health. Between 1988 and 1991, the collected data were analyzed and publications were prepared.

The aim of this epidemiological study was to collect data on healthy middle-aged men and women who are active in sports, in order to support governmental bodies in their policy on preventive (sports)-medicine. Therefore, the first goal of this study was to collect information about this specific age groups in relation to the sport activities. Secondly, this cohort was evaluated after one and two years to examine the effect of the screening. The first results are discussed in a report entitled "Effectiviteit van een medische sportkeuring bij 40-plussers, Deel 1. Maastricht, Instituut voor Bewegingswetenschappen, 1989" and further analysis of the follow-up will be published in a second report. The analysis described in this dissertation is cross-sectional, focusing on the baseline examination. The results of these analyses are useful for the interpretation of the results of the epidemiological study, and other studies on healthy and active middle-aged men and women. The design of the study and the characteristics of the selected study population will be discussed in Chapter 1. This chapter can be regarded as a general introduction to the other chapters.

One of the consequences of the approach of this epidemiological study was that the data were collected in different centers. The variability and reliability of the measurements were studied regularly during the total study period. These studies are described in Chapter 2 (flexibility measurements) and in Chapter 3 (pulmonary function tests and maximal exercise testing). Another consequence of the design of the study was that the study population was made up of an active and healthy group of middle-aged subjects. Reference values for such a population with respect to health and exercise-related parameters are hardly available.

Comparisons with reference values of less active populations and recommendations for new regression equations of pulmonary function tests are presented in Chapter 4. Reference values related to maximal aerobic exercise testing on a cycle ergometer are discussed in Chapter 6.

The criteria used for the judgement of maximal effort are discussed in Chapter 5, with special attention for the gender difference in peak plasma lactate concentration. The relationship between maximal power output on a cycle ergometer, and variables measured during a sports-medical screening such as age, anthropometric parameters, physical activity parameters, smoking habits, vital capacity and heart rate is studied in Chapter 6. Chapter 7 deals with inter-relationships between physical activity, cardiovascular fitness (maximal power output) and some cardiovascular disease risk factors. The most important results of this dissertation are summarized in Chapter 8.

Chapter 1

A medical screening in men and women over 40 years of age, active in sports

by

A.M.P.M. Bovens, M.A. van Baak,
J.G.P.M. Vrencken, J.A.G. Wijnen,
and F.T.J. Verstappen

Abstract

The aim of this chapter was to describe findings of a sports-medical check-up in men and women older than 40 years as an introduction to the following chapters. Therefore those variables which are relevant for the interpretation of results discussed in the following chapters are described. The study population existed of volunteers who were active in sports. In total 3918 subjects (2661 men and 1257 women) underwent an extensive sports-medical check-up including a medical and physical activity related history, physical examination, laboratory tests, a maximal aerobic exercise test, and an evaluation of the test results along with health and sport related advices. It appeared that the study population was not fully representative of the same age group living in the same region. Relatively more men and younger subjects were included in the study population. They were also more active in types of sport performed with higher intensities and they spent more hours per week on sport activities. Probably owing to this fact, a higher incidence rate of sport injuries was reported compared to the representative sample. Compared to the general Dutch population of the same age the study population appeared to be relatively healthy. Implications for generalization of some of the results obtained therefore should be taken into account.

Introduction

In the Netherlands until the early eighties most sports associations obliged their members to undergo a preventive sports-medical examination. Without an approval of the Federation of Sports-medical Examination Offices athletes were not allowed to take part in official sport activities organized by the associations. In the seventies a discussion was started about the value of these examinations along with the discussion whether health is the responsibility of those who are active in sports themselves, or that government and associations are responsible. The quality of the examination protocol was criticized and the effectiveness of an examination of mainly healthy people without complaints was questioned. In order to study the effectiveness of a sports-medical check-up, an extensive cross-sectional study was performed by the Institute of Sportsmedicine Limburg in collaboration with the University of Limburg and three Centers for Public Health in the most southern part of the Netherlands. Over a period of 4 years almost 4000 men and women were

examined.

In this chapter only the results of those variables collected at the check-up will be discussed, which are important for the interpretation of the results presented in the following chapters. For this interpretation it is necessary to know more about demographic aspects, physical activity and health status of the study population. For this purpose a representative sample of middle-aged men and women, living in the same region, was studied (9). The implications of a sports-medical screening in relation to health consequences and sport participation are discussed in detail elsewhere (3).

Materials and Methods

Subjects

Volunteers, active in sports for at least one hour/week for at least 3 months in the year preceding the examination, and older than 40 years of age were recruited. Initially by publications of the sports associations, sport clubs, local papers and by posters in sport accommodations. Later most of the volunteers were made aware of the examination by mouth to mouth publicity. In total 3918 volunteers (2661 men and 1257 women) were examined between March 1984 and December 1987. 97% of these subjects were active in sports at the moment of the examination, the others were allowed to take part in the study, because they intended to join a sports club after the check-up.

For practical reasons it was not possible and not desirable to examine an extensive, representative sample. In order to compare the study population with the average population of the same age, the results of a study by Van Galen and Diederiks (9) in a representative sample of the population in the same region between the age of 40 and 65 years are used. These subjects had to fill in a questionnaire with a selection of the same items used in the check-up.

Procedure

The examinations took place at three Centers for Public Health, where similar equipment was available. The subjects were examined by one of three physicians trained under the supervision of an orthopaedic surgeon, a cardiologist, a pulmonary physiologist and an exercise physiologist.

All measurements were described in protocols. During the total study period every 8 month a group of 10 subjects was examined by all three physicians at their own laboratories. In total 48 subjects (43 men and 5 women) participated in this study on the variability and reliability of measurements. The results of this study revealed acceptable results, i.e. the variation between different locations or observers was relatively small compared to the variation between subjects, so it was concluded that data can be used for a cross-sectional population study (5, 6). The protocol of the check-up is described in detail below.

Check-up

History

The history was taken by interview with a standard questionnaire. The social-economic status of the population was recorded according to education and job classification. An individual medical history (including smoking and drinking habits) as well as a family history was taken. Detailed information was asked on physical activity. Physical activity could be recorded as occupational activity (active/inactive), the use of the bicycle for transport, and sport activities. Sport activities were coded in type of sport, quartiles/year, duration in hours/week, frequency in times/week and number of years of practice. In total, information on 3 different sports could be recorded. During the research period a study with 80 volunteers out of the study population was performed to evaluate the validity of the estimation of the sport activities. These volunteers had to keep a seven-day diary on their sport activities. No significant differences were found between results of this diary and the estimation of the sport activities by the questionnaire during the check-up. Finally also detailed information was asked on sports injuries in the year preceding the check-up.

Physical examination

A general physical examination was performed including the cardiovascular, pulmonary and central nervous system. The supine blood pressure was measured by sphygmomanometry. Systolic as well as diastolic (fifth phase) blood pressure was recorded and if it was elevated it was remeasured after at least 5 min supine rest. The lowest value was used.

Anthropometric measurements included height, weight, and percentage of body fat estimated from the sum of 4 skinfolds (supra-iliacal, subscapular, biceps and triceps) according to Durnin and Womersley (8).

Visual acuity was measured by a standard method.

An important part of the physical examination consisted of measurements of the musculoskeletal function. The measurements included length of both legs, a print of both feet to judge pes cavus or pes planus, the shape of the spine, the stature of the lower limbs (genu varum or valgum) and measurements of joint motions to judge flexibility. Except for some modifications, flexibility measurements were based upon recommendations put forward by the American Academy of Orthopaedic Surgeons (1). Flexibility was quantified by recording maximal ranges of motions in 7 different joints or combination of joints. Thirteen flexibility measurements were obtained with a standard double-armed goniometer, constructed of clear flexible plastic (Protex AG, Bern). All these ranges of motion were scored accurately to 5°.

Laboratory tests

Urine was checked on the presence of glucose and protein by a Combur-test (Boehringer-Mannheim). The hemoglobin content of blood derived from a fingerprick was measured with a Compur 1000 (Zeiss). Pulmonary function tests were performed by an electronic spirometry system (Auto-Spiror Discom 14, Chest Corporation, Tokyo). The best test out of three correctly performed tests was used for analysis. The vital capacity (VC), the forced vital capacity (FVC), the volume exhaled in the first second (FEV1.0), the ratio of this volume (as percentage of the vital capacity), and the peak expiratory flow (PEFR) were recorded.

Before a maximal cycle ergometer test was carried out, a standard 12-lead resting ECG was obtained. If no contra-indications for testing (according to the Guidelines for Exercise Testing and Prescription of the American College of Sports Medicine (2)) were present, the subjects were encouraged to perform a maximal exercise test on an electrically braked cycle ergometer (Lode, Groningen). The test was performed with a pedal frequency between 60 and 80 rpm, starting at a power output of 100 W for men and 50 W for women for 5 min. Thereafter, the power output was increased by 50 W/2.5 min until a heart rate between 140 and 150 beats/min was reached. Then the increment was 25 W/2.5 min until exhaustion. During the test the ECG was recorded. At every stage measurements of blood pressure were made. Monitoring was continued for at least 7 minutes during recovery or longer if abnormalities were present. In the third minute of recovery a blood sample was drawn from a cubital vein for measurement of the peak plasma lactate concentration (640 lactate analyzer, Roche) as indication for maximal effort.

Advices and referrals

At the end of the check-up the physician discussed the results of the check-up with the subjects. In the first place they were advised on the continuation of their sport activities. Secondly different health and/or sport related advices could be given on basis of the results. For example advices on sport activities such as warming-up, stretching exercises, taping, exercise prescription as well as on general health problems such as smoking habits, (over)weight, and hygiene could be given and were recorded. It was also possible that results of the check-up could lead to a referral to the general practitioner for further evaluation.

Data analysis

The data were analysed with standard statistical packages (SPSSX and BMDP). Results are presented as percentages or mean values \pm standard deviations. Differences between groups were tested by Student t-tests, Mann-Whitney u-tests. $p < 0.05$ was considered as statistically significant.

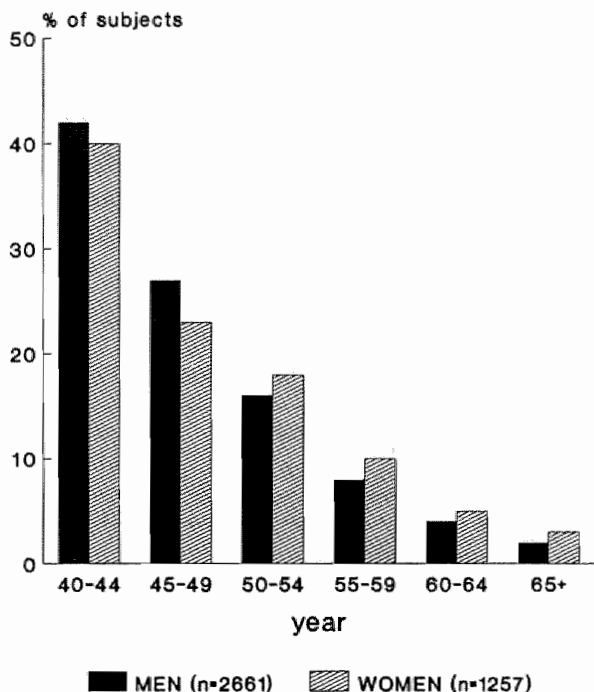
Results and Discussion

Population

Of the total population (3918 volunteers) 67.9% was male and 32.1% female ($p < 0.001$). In a representative sample of the same age group Van Galen and Diederiks (9) found that 44% of the women and 47% of the men were active in sports. Thus women were underrepresented in the present study.

In Figure 1 the distribution of age in both genders is presented. The mean age in men was 47.3 ± 6.7 years, and in women 48.1 ± 7.2 years ($p < 0.001$). A decrease of participation in sport activities with increasing age was found, as in the study of Van Galen and Diederiks (9). When both populations were compared, the present study population was younger than the active population of a representative sample. An explanation for this finding may be that older persons are not regarding themselves as athletes needing a sports-medical check-up although they are physically active in sports. This was also reflected in the difference in participation rates between types of sport. In the study of Van Galen and Diederiks (9) more recreational types of sport, such as walking, and

Figure 1:
Age distribution of the study population.



less tennis and jogging were reported (Table 2). Perhaps the same explanation can be given for the uneven distribution of gender in the present study population. About 75% of the volunteers mentioned that "a good (sports-)medical check-up" was the reason to participate in the study. Of the volunteers 39% had never had a sports-medical check-up before.

The majority of the population (87.4%) had a regular job (including house-keeping activities in women). Only 1.6% of the study population was unemployed which is considerably lower than the percentage of unemployment in the general population. Also a lower percentage of the study population was industrially disabled (5% in men and 1% in women) compared to figures of van Galen and Diederiks (9), (19% respectively 5%).

Physical activity

Physical activity was recorded as occupational activity, the use of the bicycle for transport and sport activities. More women (33.7% versus 23.6%, $p < 0.001$) reported that they were physically active in occupation. These activities in women were mainly house-keeping activities.

About half of the population used the bicycle regularly as means of transport. Significantly more women (62.9% versus 44.3%, $p < 0.001$) made use of the bicycle. An explanation for this difference may be that men make more use of the car for occupational activities outside the house compared to women and only a minority of the families have two cars.

Table 1:
Number of sports, percentage of the study population (n=3784).

	Total (%)	Men (%)	Women (%)
1 type of sport	36.0	38.1	31.5
2 types of sport	38.5	37.6	40.0
3 or more types of sport	25.6	24.3	28.5

Sport activities were recorded as types of sport, organisation form and quantity of activity (duration and frequency). In Table 1 the number of sports are presented for the study population. Maximally three types of sports were recorded. Women more often were active in more than one type of sport compared to men ($p < 0.001$). Compared to the representative sample (9) more subjects were active in more than one type of sport (64% versus 52%, $p < 0.001$).

In Table 2 numbers and percentages of the 10 most popular sports of the study population are given as well as the percentages of these sports of the representative sample (9). These figures illustrate that the study population was more active in types of sports with relatively higher intensities, like jogging and tennis. In Table 3 the percentages of the study population are presented for men and women separately. It is obvious that calisthenics and aerobics are typically female types of sport, whereas soccer is mainly played by men.

Table 2:Most popular types of sport activities of the study population and a representative sample (9)¹⁾

	Total study population (%)	Van Galen & Diederiks (%)
Jogging	31.4	10
Tennis	28.3	18
Cycling	24.4	29
Calisthenics, gymnastics	14.9	12
Swimming, waterpolo	12.3	16
Walking	11.7	32
Volleyball	11.0	7
Soccer (outdoor)	10.1	7
Soccer (indoor)	5.1	<4
Skiing	4.9	<4

¹⁾maximally 3 types of sport activities per subject could be reported

Table 3:Most popular types of sport activities of the study population (men; n = 2566 and women; n=1218)¹⁾

Men	(%)	Women	(%)
Jogging	35.7	Calisthenics	37.0
Cycling	30.4	Tennis	34.7
Tennis	25.2	Jogging	22.4
Soccer (outdoor)	14.9	Swimming, waterpolo	20.0
Volleyball	11.1	Walking	16.8
Walking	9.3	Cycling	11.9
Swimming, waterpolo	8.7	Volleyball	10.8
Soccer (indoor)	7.4	Aerobics, jazz-dancing	9.5
Badminton	5.6	Skiing	5.7
Skiing	4.6	Badminton	4.4

¹⁾maximally 3 types of sport activities per subject could be reported

About 70% of the population was active in a recreational way during sport activities compared to 83% ($p < 0.001$) in the representative sample (9) (data not shown). In Table 4 the participation in sport activities is quantified as duration (hours/week) and frequency (times/week). The total figure can be somewhat underestimated because maximally three types of sport could be recorded. Men were significantly more hours/week active than women (Table 4). Compared to the representative sample (9) the study population was active more frequently per week (2.8 times/week versus 2.3 times/week) but no difference was found in duration (both 4.0 hours/week).

Table 4:

Frequency and duration (mean \pm SD) of sport activities for the total study population ($n=3784$) and the male ($n=2566$) and female ($n=1218$) population separately

	Total	Male	Female
Frequency (times/weeks)	2.8 \pm 2.0	2.9 \pm 2.0	2.8 \pm 1.9
Duration (hours/week)	4.0 \pm 3.4	4.3 \pm 3.6	3.5 \pm 2.8 ¹⁾
¹⁾ significant difference between men and women ($p < 0.001$)			

Injury rate

The subjects were interviewed on injuries sustained during the year preceding the check-up. As definition of a sport injury, a (subjective) inconvenience that clearly was caused by or during sport activities or related to sport, was accepted. The same definition was used by others (9, 10, 11) so it was possible to make comparisons. In Table 5 the incidence rates are shown. It appeared that significantly more men were injured than women ($p < 0.001$) (Table 5). If the injury rates were expressed per 1000 hours of sport activities the incidence rate in men was 1.28 injuries/1000 hours and 1.15 injuries/1000 hours in women. More information on these injuries has been published by Vrencken et al. (11). Compared to the representative sample (9) the incidence rates were higher in the study population (person incidence rate: 21% versus 13% ($p < 0.001$) and injury incidence rate: 25% versus 16% ($p < 0.001$)). These differences can be explained by the relatively low percentage of women and older people in the study population and/or by a

different participation rate of types of sport activities in the two populations (Table 2). For a representative sample of the total Dutch population between 18 and 45 years in 1980 Kranenborg found a person incidence rate of 46% (10).

Table 5:

Sport injuries: number of injured subjects, number of injuries, and incidence rates of the total population and the male and female population separately

	Total (n=3918)	Male (n=2661)	Female (n=1257)
Number of injured subjects	827	603	224
Number of injuries	995	739	256
Person incidence rate ¹⁾	21.1%	22.7%	17.7%
Injury incidence rate ²⁾	25.4%	27.8%	20.4%
¹⁾ person incidence rate = number of injured subjects/number of subjects			
²⁾ injury incidence rate = number of injuries/number of subjects			

Health status

At the time of the check-up, 21.3% of the volunteers was under medical treatment (19.3% in men and 25.5% in women, $p < 0.001$) and 18.8% was on medication (16.2% in men and 24.3% in women, $p < 0.001$). Anti-hypertensive drugs were used most frequently (4.1% of the total population). 24.9% of the population was actual smoker (27.3% in men and 19.9% in women, $p < 0.001$). Van Galen and Diederiks (9) found higher percentages smokers (40%, 46.4% in men and 28.4% in women). The figures of this representative sample correspond well with those of the total Dutch population. Of the study population 81.9% said that they regularly consumed alcohol (87.4% in men and 70.2% in women, $p < 0.001$). Men consumed on average 149 mg alcohol/week compared to 82 mg alcohol/week in women ($p < 0.001$).

Physical examination and laboratory tests

It is clear that in this extensive screening a great variety of health related parameters was collected. Most of these variables have been discussed in

detail elsewhere (3, 13, 14) or will be analysed in future. In Table 6 the most important characteristics of the study population are given for men and women separately.

It is obvious that flexibility has to be considered an important and complex quality in sports. Analysis of the flexibility measurements indicates that flexibility is a movement specific characteristic rather than a general characteristic, meaning that the use of a general index based on one specific measurement of flexibility or a combination of measurements can not be used (12).

Table 6:

Characteristics (mean \pm SD) of the study population of measurements during the sports medical check-up for men and women separately

	Male (n=2661)	Female ¹⁾ (n=1257)
Age (years)	47.3 \pm 6.6	48.1 \pm 7.2
Education level (1-9 scale)	5.0 \pm 2.2	4.2 \pm 2.2
Smokers (%)	27.3	19.8
Regular alcohol consumption (%)	87.8	70.6
Alcohol intake (mg/week)	149 \pm 134	82 \pm 81
Height (m)	1.76 \pm 0.07	1.63 \pm 0.06
Weight (kg)	77.5 \pm 9.1	63.8 \pm 7.9
Body fat (%)	23.7 \pm 5.2	33.4 \pm 4.9
Body mass index (kg/m ²)	25.1 \pm 2.4	23.9 \pm 2.7
Supine heart rate (beats/min)	62.6 \pm 8.9	67.0 \pm 9.3
Supine systolic blood pressure (mmHg)	136 \pm 13	134 \pm 15
Supine diastolic blood pressure (mmHg)	87 \pm 8	84 \pm 8
Vital capacity (l)	5.4 \pm 0.8	3.8 \pm 0.6
¹⁾ significant difference between men and women (p<0.001) for all variables presented, tested with a Mann-Whitney u-test (age) or t-tests (other variables)		

It may be concluded that the health status of the study population was better than cross-sectional samples of the same age. For example, less obesity was found in the study population compared to the general Dutch population (7). We found that 3.3% of the men and 2.5% of the women had a body mass index > 30 (a measure of serious obesity). In a representative sample of the Dutch population between the age of 35 and 49 years these percentages were 4.2% and 5.0% respectively and in a sample between 50 and 64 years 5.4% and 10.3% (7). Also fewer ECG abnormalities suspect for coronary heart disease were found in the present male population (2.1%) compared to an asymptomatic population of Dutch male employees over the age of 35 years (5-6%) (4). We concluded that a selection of relatively young, healthy and more intensively active middle-aged persons, with males over-represented, volunteered in a preventive sports-medical check-up, and its implications for generalization of some of the results obtained should be taken into account.

References

1. American Academy of Orthopaedic Surgeons. Joint motion: Method of measuring and recording. Churchill Livingstone, Edinburgh, London and New York, 1965.
2. American College of Sports Medicine: Guidelines for exercise testing and prescription. Philadelphia PA, Lea & Febriger, 1986.
3. Van Baak M., J. Braeken, A. Bovens, J. Vrencken, J. Wijnen, F. Verstappen: Effectiviteit van een medische sportkeuring bij 40-plussers. Deel 1: De keuring; beschrijving, bevindingen en adviezen. Maastricht, Instituut voor Bewegingswetenschappen, 1989 (in Dutch).
4. Bonjer F.H.: Eindrapport van de Commissie Opsporing en Preventie van Ischaemische Hartziekten. Nederlandse Vereniging voor Arbeids- en Bedrijfs-geneeskunde, 1983 (in Dutch).
5. Bovens A.M.P.M., M.A. van Baak, J.G.P.M. Vrencken, J.A.G. Wijnen and F.T.J. Verstappen: Variability and reliability of joint measurements. *Am J Sports Med* 18: 58-63, 1990.
6. Bovens A.M.P.M., M.A. van Baak, J.G.P.M. Vrencken, J.A.G. Wijnen and F.T.J. Verstappen: Variability and reliability of pulmonary function tests and maximal exercise testing in 3 different laboratories. Chapter 3 in this dissertation.

7. Centraal Bureau voor de Statistiek: Statistische opstellen select 3, 1986.
8. Durnin J.V. and J. Womersley: Body fat assessed from total body density and its estimation from skinfold thickness. *Br J Nutr* 32: 77-97, 1974.
9. Van Galen W.C.C. and J.P.M. Diederiks: Sport bij personen van 40 tot 65 jaar in Zuid-Limburg. Maastricht, Rijksuniversiteit Limburg, 1988 (in Dutch).
10. Kranenborg N.: Sportbeoefening en blessures. *Geneeskunde en Sport* 3: 1980 (in Dutch).
11. Vrencken J.G.P.M., A.M.P.M. Bovens, J.A.G. Wijnen, J.J.H. Braeken, M.A. van Baak, L.M. Bouter and F.T.J. Verstappen. Sportbeoefening en sportblessures bij 40-plussers. *Geneeskunde en Sport* 21: 76-79, 1988 (in Dutch).
12. Vrencken J.G.P.M., M.A. van Baak, J.A.G. Wijnen, A.M.P.M. Bovens, F.T.J. Verstappen: Flexibility: A movement specific characteristic. Submitted for publication.
13. Wijnen J.A.G., A.M.P.M. Bovens, J.G.P.M. Vrencken, M.A. van Baak and F.T.J. Verstappen. Verwijzing naar de huisarts voor nader cardiologisch onderzoek op basis van een sportkeuring bij 40-plussers; resultaten na een jaar follow-up. *Geneeskunde en Sport* 22: 85-89. 1989 (in Dutch).
14. Wijnen J.A.G., M.A. van Baak, E.S. Tan, A.M.P. Bovens, J.P.G. Vrencken, F.T.J. Verstappen: Variations in exercise systolic blood pressure in physically active middle-aged men with normal and elevated blood pressure. *Int J Sports Med* 9: 412-416, 1988.

Chapter 2

Variability and reliability of joint measurements

by

A.M.P.M. Bovens, M.A. van Baak,
J.G.P.M. Vrencken, J.A.G. Wijnen
and F.T.J. Verstappen

Abstract

The purpose of this study was to determine the variability and reliability of joint measurements, as carried out by three physician observers. The intratester variation and reliability of nine different joint measurements was determined in eight healthy subjects. The measurements were taken in eight sessions by each tester. In this population also the intertester variation and reliability was determined of the three observers. This was also done in a population of middle-aged subjects, active in sports over a period of 2.5 years.

The results indicate that it is difficult to show either an improvement or worsening of a joint motion of less than 5° to 10° for most joints measured by the same tester. The intertester variation is not consistent over a longer period of time, so differences between observers during long-term studies cannot be corrected on the basis of a single study at a single point in time. The reliability of all nine joint measurements is not very high, but is probably sufficient if the results are used to compare groups within a single population and for large studies with experienced observers. Because the reliability strongly depends on the interindividual variation, it is preferable to determine the reliability for each study population.

Introduction

Although simple joint measurements with or without the use of a goniometer are widely used in a clinical setting, especially a sports medicine practice, little is known about the variability and reliability of these measurements. The variability of a measurement is a result of the intertester variation, the intratester variation, and the biological variation. All three factors are relevant in a sports medicine examination. In addition to the influence of the biological variation, the total intratester variation is also dependent on the type of education, experience, and device instruction received by the observer and the measuring errors of the goniometer. The biological variation can be reduced by standardizing the external circumstances such as warming-up, time of the day, instruction of the subject, and so on. The same factors are involved with the intertester variation. An additional factor is introduced by errors of measurement between different observers. This factor can be reduced by identical instruction for observers and using the same calibrated instru-

ments.

One of the purposes of this study was to determine the intratester variation of several goniometric measurements. The investigation was performed by three observers, which also allowed determination of the difference in intratester variation between observers. Also, the intertester variation of the observers and its development over a period of more than 2 years was studied.

The reliability as well as the variability of goniometric measurements was determined. The reliability of a measurement depends on the ratio between the real variation within a population and the total variation measured within that population. The reliability of goniometric measurements was determined separately for each observer (the intratester reliability), and for more observers (the intertester reliability).

Subjects and Methods

Experimental design

A. Intratester variation and reliability

The goniometric measurements were carried out by three physicians, each with experience in measuring joint motion, on eight healthy subjects (five males and three females), aged 30 ± 6 years. All volunteers were university personnel and familiar with the purpose of the study. Nine different goniometric measurements were performed on each subject by each observer in eight measurement sessions. These eight sessions, which took place over a period of 3 months, were standardized according to day, time, order, and instruction. After each session the results of the measurements were sealed. All nine goniometric measurements were performed by all observers within 1 hour on the same day of the week on both right and left side of the body of the subject.

B. Intertester variation and reliability

The measurements of the first session of each subject by each observer were used to determine the intertester variation and reliability. The inter-tester variation and reliability were also determined in five groups of subjects. Ten subjects were included in each group but for the last two groups only nine subjects were available. Over a period of 2.5 years these five groups of middle-aged subjects who volunteered in an extensive sports medicine examination, were examined by the same three observers

within 3 weeks at random order. In total, 48 subjects (43 males and 5 females, mean age, 48 ± 7 years) were measured. The first session of this part of the study took place at a time when the observers had little experience in measuring joint motion. The three observers were instructed identically before the study and all goniometric measurements had been taught at the same time under supervision of an experienced orthopaedic surgeon and laid down in the form of protocols.

Joint measurements

The terminology used was recommended by the American Academy of Orthopaedic Surgeons (1). All goniometric measurements were maximal joint motions, measured with a standard, double-armed goniometer (Protex AG) of 360 degrees, constructed of clear, flexible plastic, and recorded to 5°.

- Outward rotation of the shoulder

This is an active joint motion of the shoulder with both upperarms alongside the body and both elbows in 90° of flexion. The neutral starting point is in the sagittal plane. The outward rotation is recorded in degrees of motion as both forearms move outwards simultaneously in the horizontal plane with both hands in the neutral position.

- Pronation and supination of the forearm

The subject holds a stick in each hand in a vertical position. Both elbows are in 90° of flexion with both arms alongside the body. The zero starting position is the vertical position of both sticks. Pronation is the active motion of the forearm as both sticks move inward, and supination as both sticks move outward. The dorsal side of the hands and the forearm are held in the sagittal plane.

- Flexion and extension of the wrist

These are passive joint motions. The angle between the ulnar side of the forearm and the fifth metacarpal bone is measured with the zero position as hands and forearms are in the same plane. Flexion is recorded when the dorsal side of both hands are held together with the fingers pointing downward and both elbows maximally moved downward. Extension is recorded with the palms held together with the fingers pointing upward and both elbows maximally moved upward.

- Eversion and inversion of the foot

These passive joint motions are measured in supine position. The foot is aligned with the tibia in the long axis from the ankle to the knee. The examiner carries the foot through the motions of active inversion, including supination, adduction, and some degree of plantar flexion; and eversion, including pronation, abduction, and dorsiflexion. The heel must be held on the examination couch in a 90° angle between leg and foot. Eversion is recorded by turning the foot outward, inversion by turning it inward.

- Flexion and extension of the ankle

These active joint motions were also measured in a supine position. The zero starting position is with the leg at a 90° angle with the foot. Extension (plantar flexion) and flexion (dorsiflexion) are recorded.

Data analysis

For all goniometric measurements no significant difference between left and right side was found. Therefore, only results measured on the left side are reported.

- Variability

1. Intratester variation: All results are determined for each observer. For each subject the mean, standard deviation (SD), and the coefficient of variation (CV) of each goniometric measurement was determined. The CV ($SD/mean \times 100\%$) is a relative measure.

2. Intertester variation: Analyses of variance were carried out to test for differences between the three observers. The Wilcoxon matched-pairs signed-ranks test was used to compare the results of pairs of observers. $p < 0.05$ was considered as statistically significant.

- Reliability

If e represents the difference between a single observation on a subject, X , and its underlying true value, T , the classical linear model for an observed score is obtained: $X = T + e$. In a population of subjects, the error-free score T will have a variance of σ^2T and e has a variance of σ^2e . Under the assumption that the distribution of the errors is independent of the value T , the variance of X is $\sigma^2X = \sigma^2T + \sigma^2e$.

Translated into words: there are two components to the variability

among a series of measurements on different subjects. A single quantity that usefully expresses the relative magnitude of the two components of the variance of X is the reliability coefficient R :

$$R = \frac{\sigma^2 T}{\sigma^2 T + \sigma^2 e}$$

where R is the proportion of the variance of an observation due to subject-to-subject variability in error-free scores.

The intratester and intertester reliability are calculated according to Fleiss (4) and should be interpreted as follows: $R < 0.4$, poor; $0.4 < R < 0.75$, fair to good; $R > 0.75$, excellent.

Results

Intratester variation

The intratester variation as a result of the biological variation and the variation of the observation within one observer, shows for seven goniometric measurements a mean coefficient of variation for all three observers less than 10%. The CV of the eversion of the foot and the flexion of the ankle is large, not as a result of the absolute variation (standard deviation), but because the range of motion is small. If we take the mean value of the intratester CV of the three observers, we can conclude for all goniometric measurements that the mean intratester CV is less than the mean intertester CV.

Intertester variation

For all goniometric measurements, except supination and pronation, analysis of variance revealed significant intertester differences for the mean values (Table 1). Significant intertester differences in CV were only found for the supination and pronation of the forearm and the eversion of the foot (Table 1). Table 2 shows also that in groups of middle-aged subjects, active in sports, no significant intertester differences for flexion and extension of the ankle were found at most sessions. It is remarkable that intertester differences vary with time and show no consistent pattern. Even though significant differences between a pair of observers was found repeatedly, these differences were not always in the

Table 1:
Intratester and intertester variability and reliability¹⁾

	Observer				Analysis of variance
	1 (intra)	2 (intra)	3 (intra)	total (inter) ²⁾	
Outward rotation of the shoulder:					
mean±SD(°)	67.0±10.1	64.5±14.6	57.6±13.0	63.0±11.7	***
mean SD±SD(°)	5.0±1.5	6.6±2.5	5.6±2.1	7.4±2.1	ns
mean CV±SD(%)	7.7±2.8	10.7±4.1	10.2±4.2	12.3±4.0	ns
R	0.76	0.80	0.83	0.63	
Pronation of the forearm:					
mean±SD(°)	79.8±9.9	82.7±9.0	80.2±13.2	80.9±6.6	ns
mean SD±SD(°)	6.4±1.8	6.4±4.1	11.4±4.3	8.8±2.5	**
mean CV±SD(%)	8.0±1.9	7.8±4.7	14.5±6.1	11.0±3.6	**
R	0.60	0.36	0.19	0.29	
Supination of the forearm:					
mean±SD(°)	120.9±16.3	120.1±7.4	123.0±11.1	121.4±13.8	ns
mean SD±SD(°)	6.9±1.8	8.8±3.2	5.7±2.0	7.8±2.1	**
mean CV±SD(%)	5.8±1.4	7.4±3.0	4.7±1.8	6.5±1.8	**
R	0.82	0.74	0.75	0.74	
Extension of the wrist:					
mean±SD(°)	90.9±4.3	91.1±5.3	97.5±4.5	93.2±2.5	***
mean SD±SD(°)	3.2±2.9	4.3±2.1	2.8±0.7	5.1±1.3	ns
mean CV±SD(%)	3.5±3.2	4.7±2.3	2.9±0.7	5.4±1.4	ns
R	0.09	0.28	0.40	0.09	
Flexion of the wrist:					
mean±SD(°)	92.3±5.8	94.8±7.9	99.3±11.0	95.5±7.0	***
mean SD±SD(°)	4.3±2.8	4.8±2.4	3.9±1.3	5.7±2.6	ns
mean CV±SD(%)	4.5±3.1	4.9±2.3	3.5±1.8	5.9±2.4	ns
R	0.27	0.62	0.87	0.59	
Eversion in the foot:					
mean±SD(°)	26.8±8.3	28.9±5.7	28.1±6.1	27.9±5.5	*
mean SD±SD(°)	5.6±1.5	3.2±1.2	3.2±1.0	4.4±0.8	***
mean CV±SD(%)	21.6±5.9	11.2±4.6	11.7±3.8	16.3±3.3	**
R	0.57	0.66	0.75	0.56	
Inversion of the foot:					
mean±SD(°)	59.6±14.7	58.0±16.2	61.6±11.5	59.8±13.9	***
mean SD±SD(°)	4.4±1.0	4.4±1.9	5.0±1.9	6.0±2.8	ns
mean CV±SD(%)	7.6±2.4	9.2±3.6	8.5±4.3	10.4±2.6	ns
R	0.93	0.91	0.83	0.80	
Extension of the ankle:					
mean±SD(°)	69.0±10.0	70.9±9.5	69.6±10.1	69.9±9.4	*
mean SD±SD(°)	3.7±0.9	3.9±1.9	3.5±2.0	4.4±1.0	ns
mean CV±SD(%)	5.3±1.5	5.5±2.7	5.4±3.5	6.5±2.1	ns
R	0.87	0.82	0.86	0.76	
Flexion of the ankle:					
mean±SD(°)	8.0±4.9	10.5±4.9	6.2±4.5	8.3±4.2	***
mean SD±SD(°)	2.4±1.1	3.0±0.8	2.1±1.4	3.3±0.6	ns
mean CV±SD(%)	43.8±43.9	32.3±14.2	13.9±10.5	50.8±28.8	ns
R	0.75	0.64	0.74	0.63	

¹⁾Measurements taken in eight sessions on eight subjects

²⁾Mean values of all observers, the intertester reliability coefficient (R) is determined of the measurements of the first session of all subjects

*p<0.05, **p<0.01, ***p<0.001

same direction. For example, in the case of the outward rotation of the shoulder, in Session 2 and Session 4 there was a significant difference between Observer 3 and each of the other two observers, but Observer 3 measured the highest values in Session 2 and the lowest in Session 4.

Intratester reliability

For each observer, two of the nine joint motions have a low reliability coefficient ($R < 0.4$) (Table 1). It is remarkable that these are all measurements of the forearm and wrist. The extension of the wrist is an unreliable measurement for all observers, although the mean SDs and the mean CVs are low. The relatively small interindividual variation explains these low reliability coefficients. Thus, measurements with a low CV need not always be reliable measurements. The opposite can occur also, as we found with the flexion of the ankle that had fairly good reliability coefficients, but high CVs. The goniometric measurements outward rotation of the shoulder, inversion of the foot, and extension of the ankle have an excellent intratester reliability for each observer, as shown in Table 1.

Intertester reliability

Two measurements show an excellent intertester reliability (inversion of the foot and extension of the ankle) (Table 1). In this population there are two measurements with a low intertester reliability (pronation of the forearm and extension of the wrist) (Table 1). Comparing Table 1 with Table 2 shows that the intertester reliability varies between different populations as a result of differences in the interindividual variation in the observed population. Furthermore it appears from Table 2 that all goniometric measurements, except for outward rotation of the shoulder, become more reliable as time goes by. After a first session when eight of the nine goniometric measurements had poor reliability coefficients, the reliability increased, varying from fairly good to excellent. Assuming that after a year of daily measuring joint motions a certain experience had developed, we determined the reliability coefficients of the last three sessions combined (Table 2). As shown in Table 2, for a middle-aged active population the reliability of the pronation of the forearm and the extension of the wrist seems to be fair to good (0.55 and 0.58, respectively), in contrast to the results in Table 1. This applies also for all other joint motions.

		Observer				Wilcoxon				
		1	2	3	AV	1-2	1-3	2-3	R	
Outward rotation of the shoulder:										
n=10	1	47.5±14.6	50.0±13.3	43.5±11.6	ns				0.64	
n=10	2	61.5±11.3	66.0±12.4	50.5±12.3	*	ns	**	**	0.53	
n=10	3	59.0±16.5	51.0±16.8	63.5±10.0	**	*	ns	**	0.66	
n=9	4	50.0±13.2	43.9±9.3	56.1±11.1	*	ns	*	*	0.39	
n=9	5	53.9±8.2	46.7±10.0	58.9±8.2	*	ns	ns	*	0.22	
n=28	reliability coefficient of last 3 sessions								0.52	
Pronation of the forearm:										
n=10	1	119.0±15.2	104.5±13.0	92.0±11.8	**	ns	**	ns	0.04	
n=10	2	95.0±15.1	93.0±4.8	87.0±9.1	ns				0.37	
n=10	3	87.5±5.4	85.0±6.7	88.0±8.2	ns				0.47	
n=9	4	84.4±5.8	81.1±17.7	76.7±15.8	ns				0.50	
n=9	5	84.4±8.8	85.6±12.6	83.3±7.5	ns				0.67	
n=28	reliability coefficient of last 3 sessions								0.55	
Supination of the forearm:										
n=10	1	85.0±17.0	121.5±24.2	108.5±11.8	**	**	**	ns	0.06	
n=10	2	103.0±12.5	109.5±11.2	110.5±13.6	ns				0.30	
n=10	3	130.0±14.1	117.5±19.6	114.5±12.3	*	*	**	ns	0.58	
n=9	4	100.6±38.2	91.1±29.5	97.2±30.3	ns				0.89	
n=9	5	113.3±12.0	110.0±13.9	107.8±14.6	ns				0.72	
n=28	reliability coefficient of last 3 sessions								0.83	
Extension of the wrist:										
n=10	1	89.5±13.2	96.5±12.5	85.0±6.2	**	ns	ns	*	0.16	
n=10	2	89.0±7.4	95.0±8.8	83.0±9.5	*	*	ns	*	0.36	
n=10	3	92.0±4.2	92.5±6.3	89.5±1.6	ns				0.26	
n=9	4	85.6±5.8	93.3±6.6	88.3±3.5	*	*	ns	*	0.49	
n=9	5	82.8±7.1	87.8±10.3	89.4±8.5	*	*	*	ns	0.74	
n=28	reliability coefficient of last 3 sessions								0.58	
Flexion of the wrist:										
n=10	1	93.5±5.8	103.0±9.8	85.5±6.8	*	*	*	**	0.22	
n=10	2	94.0±5.2	90.5±10.1	87.5±6.3	ns				0.40	
n=10	3	99.0±8.8	96.0±7.0	90.5±1.6	*	ns	*	*	0.32	
n=9	4	89.4±25.2	86.1±15.6	83.3±13.5	ns				0.84	
n=9	5	94.4±8.1	88.3±9.0	86.7±6.1	*	ns	*	ns	0.54	
n=28	reliability coefficient of last 3 sessions								0.74	
Eversion of the foot:										
n=10	1	20.5±4.4	17.5±4.9	17.5±4.9	ns				0.06	
n=10	2	20.5±3.7	15.0±4.1	16.5±4.7	*	*	*	ns	0.24	
n=10	3	24.5±5.5	21.0±6.6	20.5±5.0	ns				0.27	
n=9	4	22.8±4.4	18.3±7.1	22.2±7.5	*	ns	ns	*	0.58	
n=9	5	23.3±4.3	20.0±4.3	18.3±4.3	*	*	*	ns	0.60	
n=28	reliability coefficient of last 3 sessions								0.45	

		Observer			AV	Wilcoxon			R
		1	2	3		1-2	1-3	2-3	
Inversion of the foot:									
n=10	1	56.0±7.7	55.5±4.4	39.5±1.4	*	ns	*	ns	0.17
n=10	2	54.5±14.8	59.5±11.1	38.5±8.5	**	ns	**	**	0.39
n=10	3	66.5±9.1	69.5±8.0	53.0±8.2	**	ns	**	**	0.31
n=9	4	56.7±10.9	66.7±12.2	62.8±8.7	ns				0.61
n=9	5	57.2±10.6	58.9±9.3	58.3±10.0	ns				0.83
n=28	reliability coefficient of last 3 sessions								0.49
Extension of the ankle:									
n=10	1	58.5±5.3	58.0±7.1	59.0±12.9	ns				0.13
n=10	2	58.0±4.8	53.5±7.1	47.0±5.9	**	*	**	*	0.18
n=10	3	64.0±4.6	59.5±5.0	59.5±5.5	ns				0.37
n=9	4	57.0±12.8	57.2±10.0	53.9±8.9	ns				0.53
n=9	5	54.4±6.3	52.8±6.2	52.2±5.1	ns				0.54
n=28	reliability coefficient of last 3 sessions								0.49
Flexion of the ankle:									
n=10	1	9.4±5.6	9.4±4.2	6.2±6.4	ns				0.21
n=10	2	10.5±5.0	9.5±4.4	7.0±4.2	ns				0.71
n=10	3	7.5±3.5	7.0±4.2	5.0±5.3	ns				0.56
n=9	4	9.4±4.6	4.4±3.0	5.0±3.5	*	*	*	ns	0.22
n=9	5	10.6±5.8	3.3±5.0	8.3±5.6	**	**	ns	*	0.69
n=28	reliability coefficient of last 3 sessions								0.57
*p<0.05, **p<0.01									

Discussion

Those who routinely take goniometric measurements should be concerned about the variability and reliability of their findings. Although since the 1920's publications about goniometric measurements have appeared, it is not clear whether these measurements yield reliable results. Most studies on goniometric measurements discussed only the introduction of specific methods of testing; they formulated normal values of several joint motions and stressed the importance of goniometric measurements for diagnosis and evaluation of orthopedic abnormalities and rehabilitation. There are few publications dealing with the variability and reliability of goniometric measurements and often the terms used were ill-defined.

As early as 1928 Hewitt (5) concluded that active wrist motions could not be reproduced accurately. Most of the early investigators were interested in the variability of joint motions of the upper extremity. Low (6) found, by measurements of flexion of the elbow and extension of the wrist, that the intratester variation was less than the intertester variation. Boone et al. (2) and our study confirmed this. Boone et al. (2) also found that the variability of joint motion was less for the upper extremity than for the lower extremity. When looking at the standard deviation,

the opposite was found in our study. For the coefficient of variation the same conclusion can be drawn. Our study shows that the intratester variation, expressed as the coefficient of variation, strongly depends on the size of the measured angle. The intratester variation expressed as the mean standard deviation for all nine goniometric measurements is 4.8° (range 2.5° to 8.1°). This means that measurements can vary $\pm 10^\circ$ (5° to 16°) in an individual ($2 \times \text{SD}$). So it will be difficult to demonstrate either an improvement or worsening of a joint motion of less than 5° to 10° for most joints, because the error of measurement in combination with the biological variation is about this size. For the coefficient of variation this can mean a value of 60% for a small-sized joint motion, as the flexion of the ankle. Without this motion all coefficients of variation are comparable with those found by Ekstrand et al. (3) when using a comparable method of testing. Boone et al. (2) found mean intratester standard deviations of 4° . If more observers take measurements in the same subjects, the inaccuracy increases because, as we found in our study, the intertester variation is higher than the intra-tester variation. The mean intertester standard deviation increases to 5.9° . Boone et al. (2) found an increase to 4.7° .

Another finding is that when measurements are made over a longer period of time, the same observer(s) does not measure always the highest or lowest value. This was not caused by knowingly changing the method of measurement by the observer(s). Thus, it is not possible to correct results of goniometric measurements of long-term studies for intertester differences on the basis of a single study.

If the results of goniometric measurements of more observers are not used to judge individuals but, for example, to study variation within a single population, the reliability of all goniometric measurements seems to be sufficient if the observers have enough experience. This applies for the middle-aged active population but not for the other study population. In this population the pronation of the forearm and the extension of the wrist have poor reliability coefficients, although the observers have the same experience. The intertester reliability strongly depends on the inter-individual variation, thus on the constitution of the investigated subjects. This means that the reliability should be determined for different study populations.

The reliability of most goniometric measurements (except perhaps the outward rotation of the shoulder) increases as time goes by. The increasing experience, and possibly also the evaluation of the results of measurements, will improve the reliability of this kind of measurements.

Conclusions

The most important conclusions of our study are:

- All goniometric measurements have a lower intra-tester variation than an intertester variation.
- The intertester variation is not consistent over a longer period of time. Differences between observers cannot be corrected on bases of a single study.
- The error of measurement in combination with the biological variation varies between on average 5° and 10° depending on the joint motion measured.
- The reliability of most goniometric measurements increases when the observers have more experience.
- The intertester reliability depends on the interindividual variation. This means that the reliability should be determined for different study populations.

References

1. American Academy of Orthopaedic Surgeons: Joint Motion; method of measuring and recording. Chicago, A.O.S.S.M., 1965.
2. Boone D.C., Azen S.P., Lin C., Spence C., Baron C., Lee L.: Reliability of goniometric measurements. *Physical Therapy* 58: 1355-1360, 1978.
3. Ekstrand J., Wiktorsson M., Oberg B., Gillquist J.: Lower extremity goniometric measurements: a study to determine their reliability. *Arch. Phys. Med. Rehabil.* 63: 171-175, 1982.
4. Fleiss J.L.: The design and analysis of clinical experiments; in *Reliability of Measurement*. New York, John Wiley and sons, 1-32, 1986.
5. Hewitt D.: The range of active motion at the wrist of women. *J. Bone Joint Surg.* 26: 775-787, 1928.
6. Low J.L.: The reliability of joint measurement. *Physiotherapy* 62: 227-229, 1976.

Chapter 3

Variability and reliability of pulmonary function tests and maximal aerobic exercise testing in 3 different laboratories

by

A.M.P.M. Bovens, M.A. van Baak,
J.G.P.M. Vrencken, J.A.G. Wijnen
and F.T.J. Verstappen

Abstract

The purpose of this study was to determine the variability and reliability of pulmonary function tests and measurements related to maximal aerobic cycle ergometer testing as carried out at three different laboratories with similar equipment but different personnel. The measurements were taken on five groups of 9-10 middle-aged subjects active in sports who volunteered in a sports-medical examination. Each subject was tested three times within 3 weeks, the five groups were tested over a period of 2.5 years.

The results showed that at the start of the study, when the observers had little experience, variability was larger and reliability was poor for a number of measurements. Thereafter, results improved and indicated that the data collected at different laboratories can be used for a cross-sectional population study, because the variation between different locations was small compared to the variation between the subjects. In maximal exercise testing the biological variation seems to dominate in the total variation between laboratories.

Introduction

In multi-center studies results of tests collected at different locations with different equipment and observers, are pooled. The variability of measurements for a given subject is the result of several sources: the intertester variation, technical error and the biological variation of the subject. The variation resulting from the combination of biological variation and technical error is also called the intratester variation. The technical error includes the variable errors of instruments, uncontrolled environmental factors, reading errors, and instruction of the subjects, and is an important factor in intratester variation and to a greater extent in intertester variation.

In order to study the variability and reliability of measurements at three different laboratories, being part of an epidemiological study on almost 4000 middle-aged men and women, active in sports, 48 subjects were examined at all three locations. Since the whole study took place over a period of more than 2.5 years, the development of this variation could be studied over this time period; the 48 subjects were recruited in 5 groups of 9-10 subjects with an interval between the groups of about eight months. In this article variability and reliability of 3 pulmonary function

tests (vital capacity, expiratory volume in the first second and the peak flow rate), of measurements during a maximal aerobic exercise test on a cycle ergometer (maximal power output, maximal systolic blood pressure, maximal heart rate), of blood pressure at rest, and of percentage body fat are evaluated.

Material and Methods

Subjects

The subjects (43 men and 5 women), all active in sports and over the age of 40 years, volunteered in a preventive sports-medical examination. Each eight months groups of 9-10 subjects were randomly recruited after the examination for the present study. After informed consent they were examined according to the same protocol at both other laboratories within 3 weeks in random order. The subjects were tested in five groups of 9-10 subjects over a period of 2.5 years. They were all Caucasians and had no signs or symptoms of diseases at the time of testing. Their characteristics are shown in Table 1.

Table 1:

Characteristics of the study population (5 groups (M = male and F = female) and total).

Group	Number	Age (y)	Height (m)	Weight (kg)
1 (10M)	n=10	44.5±4.2	1.74±0.07	77.5±11.3
2 (9M,1F)	n=10	47.4±5.1	1.73±0.10	74.8±11.6
3 (7M,3F)	n=10	43.1±2.8	1.73±0.09	73.5±7.9
4 (8M,1F)	n=9	51.0±8.4	1.72±0.08	73.6±6.4
5 (9M)	n=9	54.4±8.7	1.77±0.10	76.7±12.0
Total	n=48	47.8±7.2	1.74±0.09	75.2±9.8

General procedure

In each laboratory the tests were performed by a physician assisted by a technician. The observers were all identically instructed and were trained together. All measurements were performed according to written proto-

cols. At the three laboratories similar equipment was used. The tests took place between 9.00 a.m. and 16.00 p.m.

The percentage of body fat (BF) was estimated from the sum of 4 skinfolds (subscapular, supra-iliacal, biceps, and triceps) determined with a skinfold caliper (Servier) according to Durnin and Womersley (4). Supine blood pressure was measured by sphygmomanometry. Systolic (SBP) as well as diastolic blood pressure (DBP) were recorded. If it was elevated, it was remeasured after at least another 5 min supine rest. The lowest value was taken. Fifth-phase diastolic pressure was used.

Pulmonary function tests

Before the examination the subjects were not allowed to smoke for at least 3 hours. The electronic spirometry system (Autospiror Discom 14, Chest Corporation, Tokyo, Japan) used at the three centers detects respiration flow and volume by means of a pneumotachometer equipped with a pressure differential transducer. Volumes can be measured up to 10 liters and are automatically corrected from ambient conditions to BTPS.

The subjects were tested in a seated position and the best test out of three correctly performed tests was used for analysis. Regular checks, using a 3-liter positive displacement syringe to generate a known volume change, confirmed that the three systems produced accurate results within 3% full scale.

In the present study we describe three pulmonary function tests. The vital capacity (VC) is the largest volume measured on a complete expiration after the deepest inspiration without forced or rapid effort. The other tests were obtained from dynamic spirometry and the best out of three performances was recorded. FEV_{1.0} is the volume exhaled in the first second after maximal inspiration, as forcefully and rapidly as possible, and peak expiratory flow rate (PEFR) is the highest flow rate attained at any time during a maximal forced volume expiration manoeuvre.

Exercise test

The exercise test was performed on an electrically braked cycle ergometer (Lode, The Netherlands) with the subjects seated upright. The pedal frequency was kept between 60 and 80 rpm. Men started for 5 minutes at a power output of 100 W and women at 50 W, as warming-up.

Thereafter, the power output was increased by 50 W/2.5 min until a heart rate between 140 and 150 beats/min was reached. Then, the increment was 25 W/2.5 min until exhaustion. Maximal power output (W_{\max}) was the highest power output that could be maintained for 2.5 min. If the final power output was not completed, W_{\max} was calculated out of the completed time. Heart rate was determined from a 3-lead (II, V1 and V5) ECG, the highest rate (beats/min) was recorded as maximal heart rate (HR_{\max}). Maximal systolic blood pressure (SBP_{\max}) was recorded if possible during the final power output.

Data analysis

Variability

For each group of subjects the means and standard deviations were calculated for each observer. A two-way analysis of variance was carried out to test for systematic differences between the three laboratories. The Wilcoxon matched-pairs signed-rank test was used to compare the results of pairs of laboratories. $p < 0.05$ was considered as statistically significant. The coefficient of variation ($CV = SD/\text{mean} \times 100\%$) was determined for each subject and presented as mean and standard deviations for the five groups as well as for the total group.

Reliability

If e represents the difference between a single observation on a subject, X , and its underlying true value, T , the classical linear model for an observed score is obtained: $X = T + e$. In a population of subjects, the error free score T will have a variance of σ^2T and e has a variance of σ^2e . Under the assumption that the distribution of the errors is independent of the value T , the variance of X is $\sigma^2X = \sigma^2T + \sigma^2e$. Translated into words: there are two components to the variability among a series of measurements on different subjects. A single quantity that usefully expresses the relative magnitude of the two components of variance of X is the reliability coefficient R :

$$R = \frac{\sigma^2T}{\sigma^2T + \sigma^2e}$$

where R is the proportion of the variance of an observation due to subject to subject variability in error-free scores.

The intertester reliability coefficient was calculated according to Fleiss

(5) and should be interpreted as follows: $R < 0.4$, poor; $0.4 < R < 0.75$, fair to good; and $R > 0.75$, excellent.

Results

Table 2:

Intertester variability and reliability of body fat and blood pressure measurements in five groups (mean values \pm SD, analysis of variance (AV), mean coefficients of variation and SD (CV) and reliability coefficients (R)).

		Observer								
		1		2		3		AV	CV	R
			1-2		2-3		1-3			
Body fat (%):										
n=10	1	23.6±3.8	**	21.1±3.9	ns	21.9±4.3	**	**	8.7±3.4	.83
n=9	2	21.9±2.8		21.6±3.4		20.3±2.9		ns	6.6±3.8	.74
n=10	3	29.0±6.5	*	27.0±5.4	*	24.3±5.9	**	***	9.5±5.4	.83
n=9	4	25.6±5.3		22.9±4.2		24.1±5.4		ns	8.6±4.8	.80
n=8	5	23.2±5.7		24.6±6.4		22.4±7.5		ns	10.4±4.6	.84
n=46	tot.	24.8±5.4	**	23.4±5.0	**	22.4±5.4	***	***	8.7±4.7	.83
Systolic blood pressure (mmHg):										
n=10	1	141.5±15.3		136.9±18.3		127.0±14.0		ns	8.5±3.4	.54
n=10	2	138.0±7.9		140.0±14.1		144.5±13.6		ns	5.9±2.1	.50
n=10	3	133.5±12.9		139.0±12.0		136.0±13.1		ns	5.2±3.1	.59
n=9	4	139.4±14.5	*	147.8±19.1	ns	142.1±15.4	ns	*	5.4±2.5	.76
n=9	5	140.0±13.2		143.9±16.0		145.8±20.1		ns	4.6±2.6	.77
n=48	tot.	138.4±12.7		141.3±15.8		138.9±16.3		ns	5.9±3.0	.62
Diastolic blood pressure (mmHg):										
n=10	1	90.5±9.3		87.0±8.2		88.5±8.2		ns	3.7±1.9	.83
n=10	2	90.0±5.8		89.0±9.1		89.0±5.7		ns	3.6±2.5	.68
n=10	3	87.5±5.9		84.5±6.9		87.5±11.1		ns	4.8±4.9	.59
n=9	4	87.8±12.3		87.8±12.0		89.4±10.4		ns	4.7±3.1	.85
n=9	5	90.6±10.4		87.8±8.3		92.1±11.7		ns	3.3±2.4	.87
n=48	tot.	89.3±8.7	*	87.2±8.7	*	89.3±9.4	ns	*	4.0±3.1	.78
*p<0.05, **p<0.01, ***p<0.001; Pairs of observers (1-2, 2-3 and 1-3) are tested if significant difference is found between 3 observers (AV).										

In Table 2 the results of measurements of body fat and blood pressure are presented. For body fat in two of the five groups significant differences are found between the three observers. No systematic variation

between pairs of observers over a longer period is found. The reliability coefficients are fairly constant over 2.5 years resulting in an overall reliability of 0.83 in the total population.

Table 3:

Intertester variability and reliability of 3 pulmonary function tests (vital capacity (VC), forced expired volume 1-sec (FEV1.0) and peak flow rate (PEFR)).

		Observer			AV	CV	R
		1	2	3			
		1-2	2-3	1-3			
VC (l):							
n=8	1	5.5±0.3	5.6±0.8	5.7±0.7	ns	7.9±6.7	.79
n=10	2	5.5±1.7	5.4±1.7	5.5±1.8	ns	3.9±3.5	.99
n=9	3	5.3±1.2	5.0±1.2	5.1±1.3	ns	5.4±4.4	.93
n=9	4	5.5±1.0	5.2±1.1	5.3±0.9	ns	5.2±4.1	.92
n=8	5	5.1±0.9	5.0±0.7	5.2±0.9	ns	5.3±3.0	.87
n=44	tot.	5.4±1.1	5.3±1.1	5.4±1.2	ns	5.5±4.4	.91
FEV1.0 (l):							
n=9	1	3.8±0.7	* 4.1±0.8	ns 4.0±0.8	* *	6.2±3.9	.86
n=10	2	4.0±1.1	4.2±1.2	4.3±1.3	ns	6.0±5.1	.93
n=9	3	3.6±1.1	3.5±1.0	3.8±1.2	ns	4.4±2.9	.97
n=9	4	4.0±9.0	* 3.7±0.8	* 4.0±0.9	ns *	5.4±2.2	.96
n=9	5	3.5±0.7	3.6±0.7	3.7±0.8	ns	5.6±2.5	.92
n=46	tot.	3.8±0.9	ns 3.8±0.9	** 4.0±1.0	** ***	5.7±3.4	.93
PEFR (l/min):							
n=7	1	8.2±1.9	11.0±2.4	10.6±3.5	ns	19.1±12.2	.54
n=10	2	11.3±3.3	11.0±2.6	11.4±2.8	ns	10.1±5.5	.81
n=9	3	8.9±2.6	9.1±2.4	9.2±2.2	ns	7.7±5.2	.89
n=9	4	11.2±2.9	11.5±2.7	10.8±2.7	ns	6.4±3.8	.91
n=9	5	10.3±3.2	10.5±2.4	10.9±2.8	ns	11.0±8.5	.75
n=44	tot.	10.1±3.0	10.6±2.5	10.6±2.8	ns	10.5±8.1	.77
*p<0.05, **p<0.01, ***p<0.001; Pairs of observers (1-2, 2-3 and 1-3) are tested if significant difference is found between 3 observers (AV).							

Measurements of blood pressure showed only in one occasion significant differences between observers (SBP, group 4, Table 2). The diastolic blood pressure was less variable and had a higher reliability than systolic blood pressure at rest. Over a longer period the mean values do not indicate systematic patterns of difference between pairs of observers.

In the first group of subjects, tested at a time when the observers still

had little experience, the pulmonary function tests were most variable and least reliable (Table 3). In the later groups good reliability ($R > 0.75$) was found for all pulmonary function tests.

Table 4:

Intertester variability and reliability of measurements related to maximal exercise testing on a cycle ergometer.

		Observer			AV	CV	R
		1	2	3			
		1-2	2-3		1-3		
Maximal power output (W):							
n=10	1	251±36	246±26	246±37	ns	6.1±5.7	.73
n=10	2	236±61	228±55	227±61	ns	3.7±2.1	.98
n=9	3	242±64	231±56	230±58	ns	5.4±3.9	.96
n=9	4	246±68 *	238±63 ns	234±65 *	*	3.5±1.8	.98
n=8	5	263±75	256±62	255±64	ns	4.4±2.2	.97
n=47	tot.	250±64 **	243±56 *	240±60 ***	***	4.6±3.6	.95
Maximal systolic blood pressure (mmHg):							
n=9	1	200±15	209±13	202±21	ns	4.2±2.9	.64
n=5	2	201±15	197±30	188±24	ns	6.2±3.9	.65
n=9	3	217±15	215±12	196±9	ns	7.0±2.7	.13
n=9	4	197±14	203±15	210±17	ns	5.6±2.3	.45
n=9	5	209±14	207±25	209±31	ns	6.0±3.2	.65
n=41	tot.	205±16	207±19	202±22	ns	5.8±2.9	.51
Maximal heart rate (beats/min):							
n=9	1	170±15	164±13	169±16	ns	3.2±1.8	.44
n=9	2	178±10	177±9	179±11	ns	1.9±1.0	.79
n=9	3	174±5	175±8	178±12	ns	2.6±1.0	.71
n=9	4	168±14	168±15	170±18	ns	2.2±0.9	.94
n=8	5	167±19	165±18	162±23	ns	2.4±1.2	.83
n=44	tot.	171±14	170±14	171±17	ns	2.5±1.3	.76
*p<0.05, **p<0.01, ***p<0.001; Pairs of observers (1-2, 2-3 and 1-3) are tested if significant difference is found between 3 observers (AV).							

Maximal power output showed only significant differences between laboratories in one group (Table 4). The reliability coefficients were all excellent with an overall R of 0.95. The systolic blood pressure at maximal power output had a variation which was similar to the variation in systolic blood pressure at rest (5.8% versus 5.9%). The reliability was fair to good for 4 of 5 groups. Starting problems were also apparent for

maximal power output and maximal heart rate, which had the highest variability and lowest reliability in the first group.

Discussion

The purpose of this study was to investigate if the test results collected at three different laboratories can be used for a population study. When subjects are tested at different locations an increased variability and lower reliability can be expected. It depends on the diagnostic purpose of the tests whether this variation is acceptable. For example, when evaluating the effect of medication on pulmonary function in a patient, the reproducibility needs to be high. For this purpose tests in different laboratories may not be acceptable.

At the start of the study, when the observers had little experience, the results of variability and reliability were poor for the pulmonary function tests. Two parameters of the exercise test (maximal power output and maximal heart rate) also showed the largest variation at the very beginning of the study. After the first series and perhaps because of the evaluation thereof, the results indicate that data on the studied variables collected at different laboratories can be used for a cross-sectional population study. Although differences between laboratories existed, it can be stated that the variation because of these differences is relatively small compared to the variation between subjects.

The intertester variation is the combined result of the different intrates-ter variations and the variation between the observers. In this study it is not possible to discriminate between factors of variation because the subjects were tested only once by the same observer. No systematic differences between the three observers over the total study period were found, so it is not possible to correct for differences between observers.

When we look at the percentages of body fat it appears that results of the study population vary about 2% of body fat (mean CV = 8.7%) on average between observers. Womersley and Durnin (13) studied the effect of the observer on the estimation of the body fat content according to the same method (sum of four skinfolds). They found maximal differences, i.e. the difference between the lowest reading by any of the three observers and the highest reading by any of the three observers, between 2% and 5% of fat for men ($n = 10$) and between 3% and 6% for women ($n = 8$). They concluded that variations in skinfold thickness due to different observers, experienced and inexperienced, were not

likely to critically influence results obtained. Since the subjects of the present study were tested within 3 weeks and were not dieting or changing their training habits, the biological variation due to change in training or nutrition status may be neglected. The difference between observers in measuring body fat is not systematic which is in agreement with results from Kemper and Pieters (10) on two different observers at different institutes.

Measurements of systolic blood pressure showed a greater variation than measurements of diastolic blood pressure, which is in agreement with the physiological responses of blood pressure and has previously been demonstrated by Armitage et al. (1). Diastolic blood pressure can be measured more reliably than systolic blood pressure by different observers. The interindividual variation in the response of systolic blood pressure during a progressive maximal bicycle ergometer test in this study population has been described by Wijnen et al. (12).

In pulmonary function testing several authors investigated intertester variability in healthy subjects (2, 7, 11). Groth et al. (7) studied the intraindividual variation and the effect of learning in lung function examinations. He concluded that the intraindividual differences between the second and first examination were very small. Factors such as the subject's sex, age, height, weight, social status, motivation, whether the examinations were conducted by the same or by different technicians and whether the interval time between the two examinations was short or long, were insignificant. He found a coefficient of variation of 4.7% for the FEV_{1.0} in a non-smoking population which is somewhat smaller than 5.5% in the present study. Chinn and Lee (1) found a coefficient of variation of 2% with subjects who were repeatedly tested by the same observer in one session. The difference between both studies and the present study may be caused by the use of three apparatus by different observers in this study and by testing at different days compared to the study of Chinn and Lee (1). The difference in variation between the lung volume indices and the expiratory flow rates are in agreement with the findings of others (2, 7). McCarthy et al. (11) showed that the within-subject variability of flow volume curves between days and/or weeks was not greater than the variability of measurements within a day. Garrard and Emmons (6) did not find day to day or diurnal variation in pulmonary volume tests either.

Within-subject variation in physiological responses to exercise has been studied by several authors. In the present study regular calibration of all cycle ergometers did not reveal important deviations over time. Although

technical error of the instruments can be important, it may be concluded that the biological variation is the most important factor contributing to the intertester variation, along with the motivation of the subjects as well as of the observers. Davies et al. (3) recommended that three preliminary tests should be performed in order to obtain reproducible and reliable data. Garrard and Emmons (6) suggested that for better conditioned subjects this probably is not necessary. They found CV's of 5.5% for maximal power outputs in cycle ergometry. They tested six subjects on 3 nonconsecutive days at the same laboratory. Kuipers et al (9) tested ten subjects once a week during a 9 to 12 month period also at one laboratory. They found a CV varying between 3.0% and 6.8% (mean 5.0%). The results of our population were better except for the first group. The total study population had a mean CV of 4.6%. Compared to results in the literature it appears that the variability of measurements related to maximal exercise determined at three different laboratories is not higher than variability of different tests at one laboratory in the same subject. These results suggests that the biological variation is the most important factor of the total variation in maximal exercise testing. Since the maximal heart rate is strongly related to the maximal power output, the same pattern of variation of HRmax can be expected. This is confirmed by the results of the HRmax with an overall CV of 2.5% compared to 3.8% in the study of Garrard and Emmons (6) and 5.5% in the study of Jones and Kane (8), using data of maximal exercise tests at 4 different laboratories. The variation of the HRmax in their study was reduced to 4.6% if the data from the laboratory which showed the greatest variation in measurements were excluded. The results of both variables of maximal exercise (Wmax and HRmax) indicate that the subjects achieved similar levels of (maximal) effort at the three laboratories.

Although it is not always possible to make measurements of systolic blood pressure at maximal exercise, the variability of systolic blood pressure is comparable with that at rest. Over the total group the reliability was somewhat less at maximal exercise.

In conclusion it can be said that after a short period of practising it is possible to obtain reliable data at three different laboratories on pulmonary function tests and measurements related to maximal aerobic exercise tests, so they can be used in a cross-sectional population study. Although significant differences between the three laboratories for the total study population were found for some variables, no systematic differences were present over the total study period of 2.5 years. The

biological variation plays probably a major role in the total variation.

References

1. Armitage P., Fox W., Rose G.A., Tinker C.M.: The variability of measurements of casual blood pressure II. Survey experience. *Clin Sci* 30: 337-344, 1966.
2. Chinn D.J., Lee W.R.: Within- and between-subject variability of indices from the closing volume and flow volume traces. *Bull Eur Physiopath Respir* 13: 789-802, 1977.
3. Davies C.T.M., Tuxworth W., Young A.M.: Physiological effects of repeated exercise. *Clin Sci* 39: 247-258, 1970.
4. Durnin J.V.G.A., Womersley J.: Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged 16 to 72 years. *Br J Nutr* 32: 77-97, 1986.
5. Fleiss J.L.: The design and analysis of clinical experiments, in *Reliability of Measurements*, New York, John Wiley and Sons: 1-32, 1986.
6. Garrard C.S., Emmons C.: The reproducibility of the respiratory responses to maximum exercise. *Respiration* 49: 94-100, 1986.
7. Groth S., Dirksen A., Dirksen H., Rossing N.: Intraindividual variation and effect of learning in lung function examinations. A population study. *Bull Eur Physiopath Respir* 22: 35-42, 1986.
8. Jones N.L., Kane J.W.: Quality control of exercise test measurements. *Med Sci Sports* 11: 368-372, 1979.
9. Kuipers H., Verstappen F.T.J., Keizer H.A., Geurten P., van Kranenburg G.: Variability of aerobic performance in the laboratory and its physiologic correlates. *Int J Sports Med* 6: 197-201, 1985.
10. Kemper H.C.G., Pieters J.J.L.: Comparative study of anthropometric measurement of the same subjects in two different institutes. *Am J Phys Anthropol* 40: 341-344, 1974.
11. McCarthy D., Craig D.B., Cherniack R.M.: Intra-individual variability in maximal expiratory flow-volume and closing volume in asymptomatic subjects. *Am Rev Respir Dis* 112: 407-411, 1975.
12. Wijnen J.A.G., van Baak M.A., Tan E.S., Bovens A.M.P., Vrencken J.G.P.,

Verstappen F.T.J.: Variations in exercise systolic blood pressure in physically active middle-aged men with normal and elevated blood pressure. *Int J Sports Med* 9: 412-416, 1988.

13. Womersley J., Durnin J.V.G.A.: An experimental study on variability of measurements of skinfold thickness on young adults. *Human Biol* 45: 281-292, 1973.

40

Chapter 4

Pulmonary function in middle-aged men and women, active in sports; normal values, relation to smoking habits and sport activities

by

A.M.P.M. Bovens, M.A. van Baak,
J.G.P.M. Vrencken, J.A.G. Wijnen
and F.T.J. Verstappen

Abstract

The pulmonary function of a population of healthy middle-aged men and women, active in sports was studied. 2379 men (47.2 ± 7.0 year (mean \pm SD) and 4.3 ± 3.1 hours/week active in sports) and 1108 women (47.9 ± 6.6 year and 3.6 ± 2.5 hours/week) volunteered in a sports-medical screening at three Centres for Public Health. The values of the pulmonary function tests (vital capacity (VC), forced vital capacity (FVC), forced expiratory volume in 1 second (FEV1.0) and the peak expiratory flow rate (PEFR)) were higher compared with predicted values of a reference population. The FEV1.0/VC ratio did not show a difference in either gender. Higher values of function tests of volumes (VC, FVC and FEV1.0) were found in subjects active in endurance sports compared to subjects active in non-endurance sports in males as well as in females. New regression equations, according to smoking habits, are recommended based on these healthy, middle-aged men and women, active in sports, in order to detect abnormal pulmonary function characteristics among active middle-aged people as early as possible.

Introduction

In sportsmedical practice pulmonary function testing is often performed routinely for screening purposes. Aim of the screening is to obtain information about physical condition of individuals and to detect impairments related to exercise-induced bronchospasm or smoking. Numerous studies have demonstrated that athletes have larger lung volumes and higher flow rates than their age-, height-, and sex-matched peers (3, 4, 7, 8, 9, 11). It is still an open question whether this is the result of training over a longer period and/or of preselection, i.e. persons with better pulmonary function are more likely to participate in sport activities (4, 11).

Athletes who loose about 30% of their original spirometric values, may still fall into an apparently "normal" category based on current regression equations which are used as normative predictors (11). Thus, in an athlete a lung function impairment may not be recognized as such. Therefore Morrow et al. (11) developed new equations for young well-trained athletes in order to detect abnormal pulmonary function characteristics among athletes as early as possible. These equations can not be extrapolated to older age groups, since Morrow et al. (11) found a

positive correlation between pulmonary function tests and age in these young adults. It is well known that, after the age of 25 years, lung function declines and a negative correlation exists between age and lung function.

The purpose of this study was to investigate the pulmonary function of a population of healthy middle-aged men and women, active in sports. The results were compared with reference values recommended in the Netherlands by Quanjer et al. (12) for this age group. Relationships between pulmonary function, smoking habits and type of sport activities were also studied.

Materials and Methods

Subjects

Men and women, all active in sport and older than 40 years volunteered in a study to evaluate the effectiveness of a sports-medical screening. As part of an extensive sportsmedical examination they performed pulmonary function tests. They were tested at three Centres for Public Health in the South of the Netherlands and belong to the Caucasian race. A detailed medical and family history was taken in a standardized manner. The physician interviewed the subjects about their smoking habits and sport activities. Based on their smoking habits subjects were classified as (actual) smokers, smokers who stopped smoking (ex-smokers) and subjects who never smoked (non-smokers). The amount of cigarettes smoked daily by the subjects was categorized into four classes: incidental smoking (1), 1 to 9 cigarettes/day (2), 10 to 25 cigarettes/day (3) and 25 or more cigarettes/day (4). Frequency (times/week), duration (hours/week) and type of sport activities performed were recorded. For this study the type of sport was classified as either endurance sport (i.e. being active in at least one of following types of sport; jogging or long distance running, cycling, rowing or ice-skating) or non-endurance sport (all other sports). Volunteers with a history or signs of chronic obstructive lung disease or under control of a physician for pulmonary complaints were excluded from the analysis. In total 2379 men and 1108 women were included.

Measurements

All measurements were performed by identically instructed and trained personnel. Before the examination the subjects were not allowed to smoke for at least 3 hours. The examinations took place between 10.00 am and 16.00 pm over a total period of three and a half years. Before the pulmonary function tests, standing height (m) without shoes was measured and body weight (kg) with the subjects in light sports wear was obtained. Body fat (%) was estimated from the sum of four skinfolds (6). After the pulmonary function tests, a maximal cycle ergometer test was performed. The test was performed on a electrically braked cycle ergometer (Lode) with a pedal frequency between 60 and 80 rpm, starting at a power output of 100 W for men and 50 W for women for 5 min. Thereafter, power output was increased by 50 W/2.5 min until a heart rate between 140 and 150 beats/min was reached. Then the increment was 25 W/2.5 min until exhaustion. W_{max} was the highest power output that could be maintained for 2.5 min. If the final power output was not completed, W_{max} was calculated out of the completed time. The highest value of the heart rate during the test was considered to be HR_{max} . In the third minute of recovery, a blood sample was drawn from a cubital vein for measurement of the peak plasma lactate concentration (PPLa) (640 Lactate Analyzer, Roche). As criterium for maximal effort a $HR_{max} \geq 220 - \text{age}$ and/or $PPLa \geq 8 \text{ mmol/l}$ (men) or $\geq 5.5 \text{ mmol/l}$ (women) was used (Chapter 5). Levels of maximal oxygen uptake (VO_{2max}) in ml/kg.min were predicted from the W_{max} according to Storer et al. (14).

Three electronic spirometry systems (Autospirom Discom 14, Chest Corporation, Tokyo, Japan) were used. This system detects accurately respiration flow and volume by means of a pneumotachometer equipped with a pressure differential transducer. Volumes up to 10 liters can be measured and are automatically transformed from ambient conditions to BTPS. The subjects were tested in a seated position and the best test out of three correctly performed tests was used for analysis. Regularly checks using a 3-liter positive displacement syringe confirmed that the systems produced accurate results within 3% full scale.

The vital capacity (VC) is the largest volume measured on a complete expiration after the deepest inspiration without forced or rapid effort. The other pulmonary function tests were obtained from forced spirometry: the forced vital capacity (FVC), which is the maximum volume of air that can be expired as forcefully and rapidly as possible after maximal

inspiration. The FEV1.0 is defined as the volume exhaled in the first second of the FVC. The peak expiratory flow rate (PEFR) is the maximum flow rate, expressed in l/s, attained at any time during a FVC manoeuvre. The FEV1.0/VC ratio is the FEV1.0 expressed as a percentage of the subject actual VC.

Data analysis

Age and height of the subjects were entered into recommended equations by Quanjer et al. (12) yielding predicted lung function values. Measured and predicted values of the study population were compared, for men and women separately. Correlation coefficients and mean differences between measured and predicted values are presented. By multiple regression analysis new regression equations were developed for the pulmonary function parameters with age and height as independent variables. Full models were tested with all predictors (age, weight, and height) included and checked for interaction or nonlinearity. New equations were also developed for subgroups of the population classified according to smoking habits. Differences between sexes and between subjects active in either endurance sports or non-endurance sports were tested by t-tests and in case of a non-normal distribution (see Figure 1a and 1b) by Mann-Whitney U-tests.

Results

Figure 1 shows that most volunteers were aged between 40 and 50 years. Although both men and women were 2.9 ± 1.9 times per week active in sports, the duration of the activities was longer in men than in women ($p < 0.001$) (Table 1). A higher percentage of men was active in endurance sports. More men than women were smokers (26.4% versus 20.6% $p < 0.001$) and male smokers smoked more cigarettes than female smokers (Table 1). Most of the men who did not smoke currently were ex-smokers whereas most of the women had never smoked at all (non-smokers). The values of the pulmonary function parameters of the non-smokers have been compared with values predicted according to Quanjer et al. (12) (Table 2). The values of this subgroup of the study population were higher, except for the FEV1.0/VC. The percentual differences between measured and predicted values of the volumes and the ratio were in the same order of magnitude in men and women.

Figure 1a: Age distribution of the male study population.

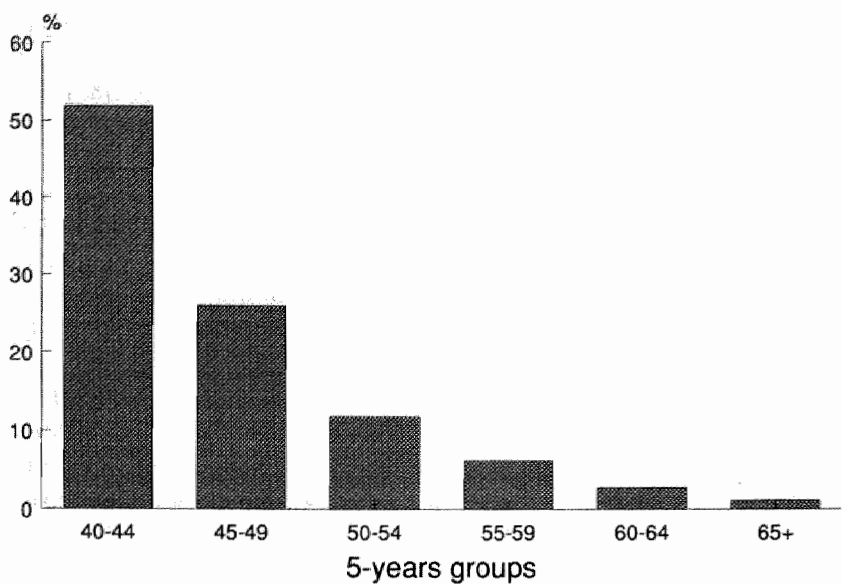


Figure 1b: Age distribution of the female study population.

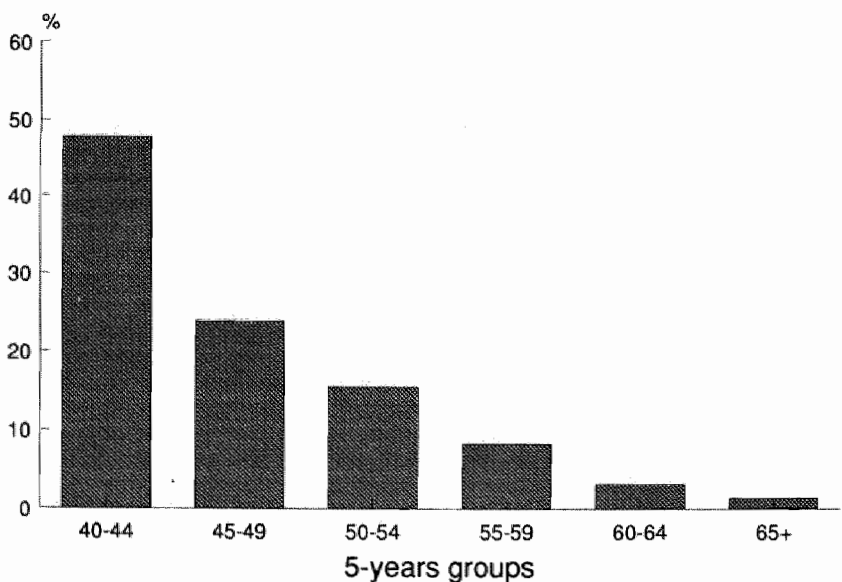


Table 1:

Characteristics of study population (means and standard deviations in 2379 men and 1108 women).

	Men	Women	
Age (year)	47.2±7.0	47.9±6.6	
Height (m)	1.76±0.07	1.63±0.06	*
Weight (kg)	77.4±8.9	63.6±7.7	*
Body fat (%)	23.5±5.2	33.2±4.9	*
Sport activities:			
duration (hours/week)	4.3±3.1	3.6±2.5	*
frequency (times/week)	2.9±1.9	2.9±1.9	ns
endurance sport (%)	61.1	46.8	*
Physical fitness:			
maximal power output (watt)	254±46	162±28	*
Wmax/kg (watt/kg)	3.3±0.7	2.6±0.5	*
VO ₂ max predicted (ml/kg.min)	41.6±6.8	29.7±4.8	*
Smoking habits:			
smokers (%)	26.4	20.6	*
ex-smokers (%)	45.6	29.1	*
non-smokers (%)	27.0	50.3	*
amount of smoking ¹⁾	1.9±1.7	1.3±1.5	
Pulmonary function tests:			
VC (l)	5.40±0.81	3.84±0.60	*
FVC (l)	5.28±0.80	3.81±0.58	*
FEV1.0 (l)	4.09±0.68	3.03±0.48	*
FEV1.0/VC (l)	77.3±7.0	79.4±6.2	*
PEFR (l/s)	10.7±2.3	7.0±1.6	*
¹⁾ Categorized into four classes: (1) incidental (2) 1-9 cigarettes/day (3) 10-25 cigarettes/day (4) 25 or more cigarettes/day (*p<0.001, ns = not significant)			

In men a higher percentual difference between measured and predicted peak flow rates was found than in women (19.3% versus 6.5%). Better correlations between measured and predicted values were found for volume variables compared to PEFR and FEV1.0/VC variables (Table 2).

Table 2:

Differences (mean, standard deviations and percentage) and correlations between actual and predicted (12) values of pulmonary function parameters for male (M; n=601) and female (F; n=520) non-smokers.

		Mean	SD	%	r
VC (l)	M	0.72	0.61	15.1	.68
	F	0.65	0.46	14.8	.66
FVC (l)	M	0.95	0.62	21.5	.67
	F	0.70	0.45	22.7	.66
FEV1.0 (l)	M	0.50	0.53	13.5	.61
	F	0.37	0.38	14.4	.63
FEV1.0/VC (%)	M	-0.25	6.75	0.3	.20
	F	-0.40	6.02	0.5	.21
PEFR (l/sec)	M	1.73	2.09	19.3	.29
	F	0.42	1.56	6.5	.34

Regression equations were calculated for each pulmonary function parameter (Table 3). The equations are given for men and women separately, as well as for subpopulations based on smoking habits. The independent variables age and height were always included in the models. Weight, a possible interaction between age and height, or exponential transformations of the independent variables did not add significantly to the prediction equations.

The subjects were further classified according to the type of sport activities. Based on analysis of the data it turned out that especially the group of subjects active in endurance types of sport showed different pulmonary function compared to the other subgroups of sport activities. Therefore, in Table 4 the results of the endurance sport group are contrasted against the non-endurance sport group for both genders. The endurance sport group reached a higher maximal power output on the cycle ergometer and spend more time on sport activities than non-endurance group. In both genders the percentages of smokers was smaller in the endurance sport group but no significant difference in the amount of cigarettes smoked by actual smokers between the subgroups were found. Pulmonary function tests of volumes (VC, FVC and FEV1.0) were significantly higher in the endurance group compared to the non-endurance group. This difference remains after correction for

smoking habits, age and height (data not shown).

Table 3:

Regression equations for middle-aged men and women, active in sports, classified according to smoking habits. General equation for all pulmonary function tests (PFT): $PFT = b_1 \cdot \text{height (m)} + b_2 \cdot \text{age (year)} + \text{constant}$.

	Sex	Height (b1)	Age (b2)	Constant	R	SEE
Non-smokers; n=601 (M), n=520 (F):						
VC	M	8.1080	-0.0258	-7.5510	.72	0.58
	F	5.8180	-0.0211	-4.6592	.66	0.45
FVC	M	7.4097	-0.0266	-6.4209	.70	0.58
	F	5.3696	-0.0245	-3.7860	.67	0.43
FEV1.0	M	4.9729	-0.0317	-3.0685	.64	0.51
	F	3.5870	-0.0272	-1.5264	.64	0.38
FEV1.0/VC	M	-15.82	-0.20	115.74	.21	6.71
	F	-18.16	-0.21	119.44	.28	5.82
PEFR	M	7.6853	-0.0561	-0.2244	.31	2.09
	F	5.0140	-0.0547	1.3119	.32	1.53
Non- and ex-smokers; n=1598 (M), n=827 (F):						
VC	M	7.3836	-0.0278	-6.2304	.68	0.59
	F	5.7586	-0.0209	-4.5514	.66	0.46
FVC	M	6.8835	-0.0297	-5.3900	.67	0.57
	F	5.1797	-0.0234	-3.5120	.66	0.43
FEV1.0	M	4.4256	-0.0342	-2.0560	.61	0.51
	F	3.5709	-0.0254	-1.5757	.63	0.37
FEV1.0/VC	M	-17.70	-0.22	-119.18	.25	6.68
	F	-14.28	-0.18	-111.80	.23	5.88
PEFR	M	7.2569	-0.0489	0.3153	.28	2.13
	F	6.4834	-0.0448	-1.4534	.33	1.53
Non-, ex- and current smokers; n=2196 (M), n=1033 (F):						
VC	M	7.2891	-0.0276	-6.0915	.67	0.58
	F	5.6446	-0.0207	-4.3836	.64	0.46
FVC	M	6.8921	-0.0298	-5.4183	.67	0.58
	F	5.1679	-0.0233	-3.5050	.65	0.43
FEV1.0	M	4.4589	-0.0338	-2.1670	.60	0.52
	F	3.5172	-0.0258	-1.4816	.62	0.37
FEV1.0/VC	M	-16.70	-0.21	116.61	.23	6.72
	F	-15.00	-0.19	113.21	.23	5.95
PEFR	M	6.9324	-0.0539	1.0404	.28	2.16
	F	6.5917	-0.0488	-1.1442	.34	1.51

Table 4:

Characteristics (mean \pm SD) of variables in men and women, active in sports, grouped according to endurance type of sport activities.

	Men		Women		
	endurance n=1356	other n=1023	endurance n=340	other n=752	
Age (year)	47.1 \pm 6.6	47.4 \pm 6.7	47.2 \pm 6.8	**	48.6 \pm 7.0
Height (m)	1.76 \pm 0.06	1.75 \pm 0.07	1.64 \pm 0.06		1.63 \pm 0.06
Weight (kg)	76.7 \pm 8.3	*** 78.3 \pm 9.5	62.5 \pm 7.8	***	64.8 \pm 7.7
Body fat (%)	23.0 \pm 5.2	*** 24.3 \pm 5.0	32.2 \pm 4.9	***	33.7 \pm 4.9
Sport activities:					
duration (hrs/wk)	5.4 \pm 3.1	*** 3.5 \pm 3.1	4.2 \pm 2.5	***	3.1 \pm 2.5
frequency (times/wk)	3.6 \pm 2.1	*** 2.0 \pm 1.4	3.5 \pm 2.0	***	2.3 \pm 1.6
Smoking habits:					
smokers (%)	21.2	*** 35.0	17.4	*	22.8
amount ¹⁾	1.9 \pm 1.7	2.0 \pm 1.7	1.3 \pm 1.5		1.3 \pm 1.5
Physical fitness:					
Wmax (W)	264 \pm 47	*** 227 \pm 42	170 \pm 29	***	151 \pm 28
Pulmonary function tests:					
VC (l)	5.46 \pm 0.8	*** 5.32 \pm 0.8	3.90 \pm 0.6	**	3.79 \pm 0.6
FVC (l)	5.34 \pm 0.8	*** 5.21 \pm 0.8	3.87 \pm 0.6	**	3.76 \pm 0.6
FEV1.0 (l)	4.12 \pm 0.7	** 4.05 \pm 0.7	3.07 \pm 0.5	**	2.99 \pm 0.5
FEV1.0/VC (%)	77.1 \pm 6.9	77.5 \pm 7.1	79.4 \pm 6.1		79.4 \pm 6.3
PEFR (l/s)	10.7 \pm 2.2	10.6 \pm 2.3	7.3 \pm 1.6	*	7.1 \pm 1.6
*p<0.05, **p<0.001, ***p<0.001, ¹⁾ categorized as in Table 1					

Discussion

The purpose of this study was to compare pulmonary function tests of middle-aged men and women, active in sports, and those of a reference population of the same age. The results indicate that differences are considerable and justify the development of new equations for healthy middle-aged men and women, active in sports. Accepting the idea that

reference values should give information about desired levels of a certain function test variable, it is logic to select a group of physically fit subjects, active in sports, as a reference group in order to make a fair comparison if active individuals or groups are tested on this particular function.

The relatively greater difference between the measured value and the predicted value for FVC than for VC indicates that the exercise-dependent influence at the first part of the FVC performance is even more pronounced in the study population. The correlation coefficients of the relations between actual and predicted values were higher for lung volumes than for the FEV_{1.0}/VC ratio and the peak flow rate. Morrow et al. (11) found also that actual and predicted values of the ratio did not correlate well. This finding may be explained by the fact that the ratio is a combined measure with its implication for comparisons. The flow rates also showed low correlations between actual and predicted values. The average residual standard deviation of the flow rates was about two times as high as the average standard deviation of the values of Quanjer et al. (12), but of about the same level as in the study of Morrow et al. (11) in young athletes. Morrow et al. (11) suggested that this was probably due to the fact that flow rates are less reproducible than volumes due to the steepness of the flow-volume curve, where small errors of measurements result in large flow differences. When the mean values of the peak flow rate are higher, the standard deviation will be proportionally greater.

The variability of mean values of pulmonary function parameters between tested populations is considerable. Quanjer et al. (12) computed equations of reference values after exploring the literature and making (graphical) comparisons of reference values. The summary equations were derived in Caucasians for an age range of 25 to 70 years and for a height range of 1.55-1.95 m in men and 1.45-1.80 m in women, respectively. Are the differences between our study population and the reference values of Quanjer et al. (12) real differences or due to methodological and technical bias? Quanjer et al. (12) suggested that different factors may account for differences in general. Firstly, ethnic differences may influence the outcome of pulmonary function tests, as also concluded by Bande et al. (2). Since in our population as well as in Quanjer's only Caucasians were included, this factor can be neglected. Different instrumentation and different methods used in the performance of pulmonary function tests may also account for a certain variability. This variation can be caused by the use of other equipment, in this case a electronic

spirometry system, compared to other reference studies. No comparative study exists where similar systems was used. On the other hand regular checks of the system with a 3-liter syringe indicated no significant variation throughout the study. Therefore it is unlikely that this factor can explain the observed differences. Other variation might have been introduced by collecting the data at three centers. This variation was minimized by instructing and training the personal together, according to the recommendations of the American Thoracic Society (1), and by using three identical measurement systems. A study on the variability and reliability of these measurements among the three different laboratories of our study revealed acceptable results (Chapter 3).

Differences in selection criteria for populations may also introduce variation. Criteria may include health, age, smoking habits, and physical fitness. As stated before, it is generally accepted that in order to obtain normal values only subjects free from signs or symptoms of pulmonary disease should be selected, which was the case in the study of Quanjer et al. (12). Therefore subjects with a history or signs of pulmonary disease were excluded from our study.

The interpretation of age-related changes of pulmonary function depends on the age distribution of the reference population. For instance, it is well known that reference values of pulmonary function do not decline with age between 18 and 25 years (10). Although we only set a minimum age of 40 years for participation in our study, most of the volunteers were between 40 and 50 years (Figure 1). Therefore, the developed regression equations reflect the decline of pulmonary function best in this age group. On the other hand it appears that in independent factor age does not vary much compared with results of Quanjer et al. (12).

Several studies indicate differences in pulmonary function between smokers and non-smokers (2, 5). Because there is a need for equations which allow to take smoking habits into account we decided to present regression equations for non-smokers, for non- and ex-smokers, as well as for the total population.

Probably the most important factor causing variability is the level of physical fitness of the populations studied. Since in our study only subjects were included who were active in sports, it may be assumed that the level of physical fitness is higher in the study population than in a general population. On the other hand, our study population should not be regarded as well-trained athletes (the majority was active in recreational sport activities). Several studies showed that athletes achieve higher pulmonary function values than non-athletes of the same age, height,

and sex (3, 4, 7, 8, 9, 11). These studies were performed in young and well-trained athletes. The present study demonstrates that middle-aged men and women, active in sports, achieve higher values too. The higher levels of lung volume in subjects active in endurance types of sport are pointing in the same direction but it must be stated that the differences between the endurance group and the "others" are considerably smaller than the differences between the "other" group and reference values. Robinson and Kjeldgaard found an improvement in ventilatory muscle function with running (13). This improvement was demonstrated in maximal expiratory pressure, maximum sustainable ventilatory capacity for 15 minutes and maximum voluntary ventilation at the end of a 20-week physical fitness program (endurance running). At the end of this 20-week period they did not find an increase in FEV_{1.0} and FVC values (13). The question remains open whether higher pulmonary function values in subjects in the present study were the result of training for several years or of preselection.

It may be concluded, that middle-aged men and women, active in sports, have higher pulmonary function values compared to reference values. The selection of an active population are likely to be responsible for these differences. New prediction equations are recommended in healthy, active and middle-aged men and women according to smoking habits. Longitudinal studies are necessary to answer the question what causes these differences.

References

1. American Thoracic Society: ATS Statement-Snowbird Workshop on Standardization of Spirometry. *Am Rev Res Dis* 119: 831-838, 1979.
2. Bande J., Clement J., Van De Woestijne K.P.: The influence of smoking habits and body weight on vital capacity and FEV₁ in male air force personnel: a longitudinal and cross-sectional analysis. *Am Rev Res Dis* 122: 781-790, 1980.
3. Bertholon J.F., Carles J., Teillac A.: Assessment of ventilatory performance of athletes using the maximal expiratory flow-volume curve. *Int J Sports Med* 7: 80-85, 1986.
4. Biersteker M.W.A., Biersteker P.A.: Vital capacity in untrained healthy young adults in the Netherlands. *Eur J Appl Physiol* 54: 46-53, 1985.

5. Dockery D.W., Speizer F.E., Ferris B.J., Ware J.H., Louis T.A., Spiro A.: Cumulative and reversible effects of lifetime smoking on simple tests of lung function in adults. *Am Rev Respir Dis* 137: 286-292, 1988.
6. Durnin J.V.G.A., Womersley J.: Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged 16 to 72 years. *Br J Nutr* 32: 77-97, 1986.
7. Folinsbee L.J., Wallace E.S., Bedi J.F., Horvath S.M.: Exercise respiratory pattern in elite cyclists and sedentary subjects. *Med Sci Sports Exerc* 15: 503-509, 1983.
8. Hagberg J.M., Yerg J.E., Seals D.R.: Pulmonary function in young and older athletes and untrained men. *J Appl Physiol* 65: 101-105, 1988.
9. Kaufman D.A., Swenson E.W., Fencel J., Lucas A.: Pulmonary function of marathon runners. *Med Sci Sports* 6: 114-117, 1974.
10. Knudson R.J., Slatin R.C., Lebowitz M.D., Burrows B.: The maximal expiratory flow-volume curve. Normal standards, variability and effects of age. *Am Rev Respir Dis* 113: 587-600, 1976.
11. Morrow J.R., Van Handel P.J., Bradley P.W.: Development of valid pulmonary function equations for trained athletes. *Int J Sports Med* 10: 43-47, 1989.
12. Quanjer Ph.H.: Standardized lung function testing. *Bull Eur Physiol Res Clin Res Physiol* 19 (suppl 5): 1983.
13. Robinson E.P., Kjeldgaard J.M.: Improvement in ventilatory muscle function with running. *J Appl Physiol* 52: 1400-1406, 1982.
14. Storer T.W., Davis J.A., Caiozzo V.J.: Accurate prediction of $\dot{V}O_{2\max}$ in cycle ergometry. *Med Sci Sports Exerc* 22: 704-712, 1990.

Chapter 5

Peak plasma lactate concentration after maximal exercise testing in people of 40 to 65 years: difference between genders

by

A.M.P.M. Bovens, M.A. van Baak,
J.G.P.M. Vrencken, J.A.G. Wijnen
and F.T.J. Verstappen

Abstract

The purpose of this study was to investigate the gender difference in peak plasma lactate concentration (PPLa) in a middle-aged population. A PPLa ≥ 8 mmol/l is generally considered as indicative for a maximal effort in a progressive (maximal) exercise test at this age. 2164 men and 975 women, all active in sports and aged between 40 and 65 year, volunteered for a medical examination including a progressive exercise test on a cycle ergometer to exhaustion. In the third minute of recovery a venous blood sample was taken to determine the PPLa. PPLa and maximal heart rate (HRmax) were lower in women than in men ($p < 0.001$). The average age of the women was slightly higher than that of the men ($p < 0.01$). Multiple regression analysis was performed to assess the contribution of gender to PPLa, independently of age and HRmax. It is concluded that PPLa is about 2.5 mmol/l lower in women than in men when age and HRmax are the same. In studies where both HRmax and PPLa are measured, we suggest to use the following criteria for maximal effort: HRmax ≥ 220 -age beats/min and/or PPLa ≥ 8 mmol/l in men respectively HRmax ≥ 220 -age beats/min and/or PPLa ≥ 5.5 mmol/l in women. In our population 88% of the men and 85% of the women met these criteria.

Introduction

Maximal exercise tests are performed widely to evaluate the aerobic power of patients or healthy persons in research as well as in sports-medical examinations. It is always difficult to judge whether a subject performed really up to his or her true maximum. There are several, more or less objective criteria that are used to judge the level of exhaustion in progressive exercise testing. It is generally accepted that maximal effort is confirmed by levelling-off of the oxygen consumption with increasing power output (2). Other parameters used are a respiratory exchange ratio above 1.00, 1.05 or 1.10, a respiratory equivalent above 30 or combinations of these and other criteria (5, 11, 12, 14).

If no gas analysis equipment is available the maximal heart rate and the peak blood lactate concentration are often used. For practical purposes maximal heart rate is often assumed to be 220 minus age (beats/min) on average, with a standard deviation of ± 10 beats/min at a given age, for men as well as women. As a criterion of maximal effort, therefore 200

minus age ($220 \text{ minus } (\text{age} + 2\text{SD})$) beats/min for maximal heart rate is used (2, 12). Åstrand (2) suggested a peak lactate concentration of 8 mmol/l or higher as a criterion of maximal effort. A possible difference between men and women was not mentioned. Jacobs et al. (9) found significantly lower muscle lactate concentrations after maximal exercise in women than in men, suggesting that there may be gender differences in peak blood lactate concentration. The aim of this study was to investigate gender differences in peak plasma lactate concentration in a healthy middle-aged population.

Subjects and Methods

Subjects

All volunteers were aged 40 years or older and active in sports (participation rate of 4.0 ± 3.4 hours per week). They performed a progressive maximal exercise test, as part of a sports-medical examination. Tests of subjects using cardiovascular medication were excluded from the analysis. Only subjects with complete data regarding maximal heart rate (HRmax), peak plasma lactate concentration (PPLa) and maximal power output (Wmax) were analysed. Data on 3139 volunteers (2164 men and 975 women) were obtained over a period of 4 years.

Table 1:
Characteristics of the study population.

	Male n=2164	Female n=975	
Age (year)	47.0 \pm 6.4	47.9 \pm 6.8	**
Weight (kg)	77.3 \pm 8.8	63.6 \pm 7.6	***
Height (m)	1.76 \pm 0.06	1.64 \pm 0.06	***
Body fat (%)	23.5 \pm 5.1	33.2 \pm 4.9	***
Wmax (W)	250.9 \pm 48.0	157.8 \pm 29.3	***
Wmax/kg (W/kg)	3.3 \pm 0.7	2.5 \pm 0.5	***
HRmax (beats/min)	173.0 \pm 13.6	168.4 \pm 13.5	***
PPLa (mmol/l)	10.5 \pm 3.2	7.7 \pm 3.7	***
p<0.01 (Mann-Whitney u-test), *p<0.001 (t-test)			

Characteristics of the study population including measurements of height, weight, and percentage of body fat, estimated from the sum of four skinfolds determined according to the method of Durnin and Womersley (4), are presented in Table 1.

Exercise test

The subjects were seated upright on a electrically braked cycle ergometer (Lode) with 3-lead (II, V1 and V5) ECG-electrodes in place. The test was performed with a pedal frequency between 60 and 80 rpm, starting at a power output of 100 watt (W) for men and 50 watt for women for 5 minutes. Thereafter, power output was increased by 50 W/2.5 min until a heart rate between 140 and 150 beats/min was reached. Then the increment was 25 W/2.5 min until exhaustion. Wmax was the highest power output that could be maintained for 2.5 min. If the final increment of power output could not be maintained for 2.5 min, Wmax was calculated with a correction for the completed time of that increment. Heart rate was determined from the ECG at the end of each power output, and the highest value recorded was considered to be HRmax. In the 3rd minute of recovery, a blood sample was drawn from a cubital vein for measurement of PPLa (640 Lactate Analyzer, Roche) (6).

Data analysis

Means and standard deviations of variables were calculated separately for men and women. Differences between men and women were tested by two-sample t-tests, except for differences in age which were tested by Mann-Whitney u-tests because of non-normal distribution. Correlation coefficients between gender, age, HRmax and PPLa were determined. Multiple linear regression analysis was used to evaluate the influence of the independent variables gender, age and HRmax on PPLa.

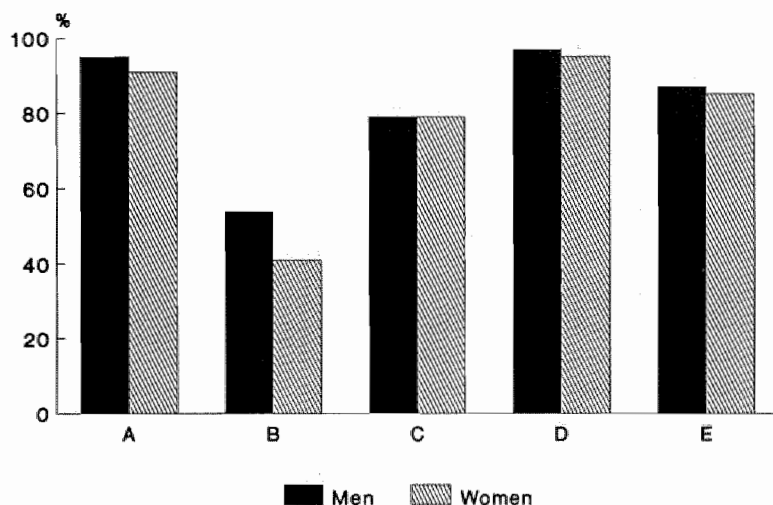
Results

A small, but significant, difference in age (0.9 year, $p < 0.01$) between the sexes was found (Table 1). Moreover, women reached significant lower HRmax than men (168.4 ± 13.5 versus 173.0 ± 13.6 beats/min, $p < 0.001$). PPLa was 2.8 mmol/l ($p < 0.001$) lower in women compared to men. Only 40% of the women reached a PPLa value of 8 mmol/l or

higher, compared to 79% of the men. To correct for a lower effort level of the women, as suggested by the lower HRmax, multiple linear regression analysis was performed with PPLa as dependent and age (year), gender (male = 1, female = 0) and HRmax (beats/min) as independent variables, ($n = 3139$): $\text{PPLa (mmol/l)} = 0.067 \times \text{HRmax} + 2.5 \times \text{gender} - 0.055 \times \text{age} - 0.91$, ($r = .48$, $p < 0.001$). This analysis shows a gender difference for PPLa of 2.5 mmol/l independent of age and HRmax.

Figure 1:

Percentages of population meeting different criteria of maximal effort based on HRmax and PPLa.



A: $\text{HRmax} \geq 200$ -age beats/min

B: $\text{HRmax} \geq 220$ -age beats/min

C: $\text{PPLa} \geq 8$ mmol/l (Men), $\text{PPLa} \geq 5.5$ mmol/l (Women)

D: $\text{HRmax} \geq 200$ -age beats/min and/or $\text{PPLa} \geq 8$ (Men) or ≥ 5.5 (Women) mmol/l

E: $\text{HRmax} \geq 220$ -age beats/min and/or $\text{PPLa} \geq 8$ (Men) or ≥ 5.5 (Women) mmol/l

Figure 1 shows the percentages of the study population meeting different criteria of maximal effort based on PPLa and HRmax. Characteristics of the subpopulations based on these criteria are presented in Table 2.

Table 2:
Characteristics of groups (A-E) based on different criteria levels.

	Sex	A	B	C	D	E
Age (year)	M	47.0±6.4	47.6±6.4	46.4±5.8	46.9±6.3	46.7±6.1
	F	47.9±6.8	48.4±7.2	47.1±6.1	47.7±6.7	47.4±6.5
Weight (kg)	M	77.2±8.7	77.0±8.3	77.0±8.5	77.2±8.7	77.1±8.5
	F	63.5±7.6	62.8±7.4	63.4±7.7	63.4±7.5	63.4±7.6
Height (m)	M	1.76±0.06	1.76±0.06	1.76±0.06	1.76±0.06	1.76±0.06
	F	1.63±0.06	1.63±0.06	1.64±0.06	1.63±0.06	1.64±0.06
Body fat (%)	M	23.5±5.1	23.7±4.9	23.4±5.0	23.5±5.1	23.4±5.0
	F	33.2±4.9	33.0±4.8	33.0±5.0	33.2±4.9	33.0±5.0
Wmax (W)	M	252±47	255±47	257±45	252±47	255±46
	F	159±29	160±29	163±28	159±29	162±28
Wmax/kg (W/kg)	M	3.3±0.6	3.3±0.6	3.4±0.6	3.3±0.6	3.3±0.6
	F	2.5±0.5	2.6±0.5	2.6±0.5	2.5±0.5	2.6±0.5
HRmax (beats/min)	M	175±12	181±9	175±12	174±13	175±12
	F	171±11	179±7	171±12	170±12	171±12
PPLa (mmol/l)	M	10.7±3.1	11.0±3.0	11.6±2.5	10.7±3.1	11.2±2.8
	F	7.9±3.7	8.3±2.7	8.6±3.6	7.9±3.7	8.4±3.6

A: HRmax \geq 200 - age beats/min

B: HRmax \geq 220 - age beats/min

C: PPLa \geq 8 mmol/l (M), PPLa \geq 5.5 mmol/l (F)

D: HRmax \geq 200 - age beats/min and/or PPLa \geq 8 (M) or \geq 5.5 (F) mmol/l

E: HRmax \geq 220 - age beats/min and/or PPLa \geq 8 (M) or \geq 5.5 (F) mmol/l

In Table 3 the correlation coefficients between age, HRmax and PPLa are presented for men who achieved a HRmax \geq 220 minus age beats/min and/or PPLa \geq 8 mmol/l and for women who achieved a HRmax \geq 220 minus age beats/min and/or PPLa \geq 5.5 mmol/l. The correlation coefficients between age and PPLa and between PPLa and HRmax are low but statistically significant.

Table 3:

Correlation coefficients between HRmax, PPLa and age in men and women who met the criterion of HRmax \geq 220 - age beats/min and/or PPLa \geq 8 (men) or \geq 5.5 (women) mmol/l (group E, Table 2).

	Male		Female	
HRmax versus age	-0.30	***	-0.37	***
PPLa versus age	-0.21	***	-0.11	***
HRmax versus PPLa (adjusted for age)	0.13	***	0.10	**
	0.08	***	0.06	*
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$				

Discussion

If no gas analysis is performed in exercise testing, HRmax and PPLa can be used as objective criteria of maximal effort. The purpose of the study was to investigate the influence of gender on PPLa after progressive maximal exercise testing. There are significant differences between men and women for PPLa (10.5 ± 3.2 respectively 7.7 ± 3.7 mmol/l) but also for HRmax (173.0 ± 13.6 respectively 168.4 ± 13.5 beats/min) in our study population. According to Hossack and Bruce (8) no such difference in HRmax between genders in this age group was expected. This suggests that the level of effort was lower in the women than in the men in our study. It is possible that this factor is responsible for the gender difference in PPLa in combination with the difference in age between the female and male group. To adjust for these factors multiple regression analysis was performed. This analysis shows that a significant gender difference of 2.5 mmol/l still remains for PPLa, independent of age and HRmax. It is important to stress that in both genders equal methods were used, so gender differences can not be the result of differences in methods used in men and women. Thus we conclude that a relevant gender difference in PPLa is present and should be taken into account in maximal exercise testing.

In the literature studies showing a difference in PPLa between men and women can be found (1, 5, 16). Froberg and Pederson (5) found a difference in maximal blood lactate concentration between men ($n = 7$) and

women ($n = 6$) of 2 mmol/l (12.3 ± 1.9 respectively 10.3 ± 1.7 mmol/l), but it was not statistically significant. Verstappen et al. (16) found a significant gender difference (25 men and 11 women) in PPLa of 2.3 mmol/l. Sidney and Shephard (15) found blood lactate levels of 11.1 mmol/l in 19 sedentary men (age, 63.7 ± 2.7 years) and 9.1 mmol/l in 20 women (age, 63.4 ± 3.6 years) making all a good effort, i.e. realizing a plateau of the oxygen consumption in a progressive treadmill test. They stated that the lower blood lactate of the females may indicate a smaller ratio of the muscle mass to the total blood volume, a factor which also contributes to the decline of maximum lactate values with aging (15). Jacobs et al. (9) mentioned significant differences in muscle lactate concentration between men and women after performing a Wingate anaerobic test. It is suggested that female skeletal muscle has a lower glycogenolytic potential reflected in lower phosphorylase and phosphofructokinase activities (10, 13). Although lactate within the active muscle is not in equilibrium with that of the extracellular space and, consequently, of blood (7), a lower lactate production rate in women is consistent with the lower PPLa concentration that we found.

From the correlation coefficients in Table 3 it appears that the correlation of PPLa with age is not very strong. Binkhorst et al. (3) found higher correlations in men but, because of methodological differences, such as the selection of the study population, it is difficult to compare these studies. One factor that may play a role is the relatively small age distribution in our population. In establishing the criteria for maximal effort in this population we decided not to include age, without suggesting that it may not be appropriate in other situations.

Besides PPLa, heart rate is also often used as a criterion of maximal effort. Heart rate is a good indicator for the stress of the cardiovascular system, whereas peak blood lactate concentration reflects more the local stress of exercising muscles. Table 3 shows that the correlation between these variables is low (also when adjusted for age), which indicates that they are largely independent physiological parameters. Thus a combined criterion for maximal effort with both HRmax and PPLa seems valid.

Several studies determined regression equations of maximal heart rate vs age (3, 8). Åstrand (2) stated that it is quite meaningless to be too sophisticated in the choice of the target heart rate during maximal exercise testing. He suggested a simple formula, assuming the maximal heart rate to be 220 minus the subject's age. The standard deviation of HRmax is ± 10 beats/min. Thus, on theoretical grounds, about half of a population is expected to achieve a HRmax ≥ 220 minus age and about

95% a $\text{HR}_{\text{max}} \geq 200$ minus age. In our population about 55% of the men and 41% of the women met this criterion of $\text{HR}_{\text{max}} \geq 220$ minus age and 95% respectively 91% met the $\text{HR}_{\text{max}} \geq 200$ minus age criterion (Figure 1). A considerable percentage of the male and female population reached a combination of the HR_{max} criterion of 220 minus age and/or PPLa (88% respectively 85%). On the other hand the results of the combination of the HR_{max} criterion of 200 minus age and/or PPLa (D, Table 2) are similar to those of the HR_{max} criterion of 200 minus age alone (A, Table 2), indicating that this HR_{max} criterion dominates the PPLa criterion.

Based on these data, we suggest the following criteria for maximal effort: $\text{HR}_{\text{max}} \geq 220$ -age beats/min and/or $\text{PPLa} \geq 8$ mmol/l in men or $\text{HR}_{\text{max}} \geq 220$ -age beats/min and/or $\text{PPLa} \geq 5.5$ mmol/l in women.

References

1. Åstrand I., Åstrand P-O., Hallböck I., Kilbom A.: Reduction in maximal oxygen uptake with age. *J Appl Physiol* 35: 649-651, 1973.
2. Åstrand P-O., Rodahl K.: *Textbook of Work Physiology*. New York, McGraw Hill, 1986.
3. Binkhorst R.A., Pool J., Leeuwen van P., Bouhuys A.: Maximum oxygen uptake in healthy nonathletic males. *Int Z angew Physiol und Arbeitsphysiol* 22: 10-18, 1966.
4. Durnin J.V., Womersley J.: Body fat assessed from total body density and its estimation from skinfold thickness. *Br J Nutr* 32: 77-97, 1974.
5. Froberg K., Pederson P.K.: Sex differences in endurance capacity and metabolic response to prolonged, heavy exercise. *Eur J Appl Physiol* 52: 446-450, 1984.
6. Geyssaut A., Dormois D., Barthelemy J.C., Lacour J.R.: Lactate determination with the lactate analyser LA 640: a critical study. *Scand J Clin Invest* 45: 145-149, 1985.
7. Gollnick P.D., Bayly W.M., Hodgson D.R.: Exercise intensity, training, diet, and lactate concentration in muscle and blood. *Med Sci Sports Exerc* 18: 334-340, 1986.
8. Hossack K.F., Bruce R.A.: Maximal cardiac function in sedentary normal men and women: comparison of age-related changes. *J Appl Physiol* 53: 799-804, 1982.

9. Jacobs I., Tesch P.A., Bar-Or O., Karlsson J., Dotan R.: Lactate in human skeletal muscle after 10 and 30 s of supramaximal exercise. *J Appl Physiol* 55: 365-367, 1983.
10. Komi P.V., Karlsson J.: Skeletal muscle fiber types, enzyme activities and physical performance in young males and females. *Acta Physiol Scand* 103: 210-218, 1978.
11. Kuipers H., Verstappen F.T.J., Keizer H.A., Geurten P., van Kranenburg G.: Variability of aerobic performances in the laboratory and its physiological correlates. *Int J Sports Med* 6: 197-201, 1985.
12. Löllgen H., Ulmer H.V.: Ergometrie - Empfehlungen zur Durchführung und Bewertung ergometrischer Untersuchungen. *Klin Wochenschr* 63: 651-677, 1985.
13. Nygaard E.: Skeletal muscle fiber characteristics in young women. *Acta Physiol Scand* 112: 299-304, 1981.
14. Saris W.H., Noordeloos A.M., Ringnalda B.E., van 't Hof M.A., Binkhorst R.A.: Reference values for aerobic power of 4-18 year-old Dutch children: Preliminary Results. In: *Children and Exercise XI*, eds. Binkhorst R.A., H.C. Kemper, and W.H. Saris, Human Kinetics Publ., Inc. Champaign: 151-160, 1985.
15. Sidney K.H., Shepard R.J.: Maximum and submaximum exercise tests in men and women in the seventh, eighth, and ninth decades of life. *J Appl Physiol* 43: 280-289, 1977.
16. Verstappen F.T., Janssen G.M., Does R.J.: Effects of endurance training and competition on exercise tests in relatively untrained people. *Int J Sports Med* 10 (suppl.): S126-S131, 1989.

Chapter 6

Maximal aerobic power in cycle ergometry in middle-aged men and women, active in sports, in relation to age and physical activity

by

A.M.P.M. Bovens, M.A. van Baak,
J.G.P.M. Vrencken, J.A.G. Wijnen,
W.H.M. Saris and F.T.J. Verstappen

Abstract

Reliable standards that may be applied for middle-aged and active men and women are desirable in sports-medical practice to allow valid comparison between measurements at maximal power output and those expected for these subjects. For this purpose maximal cycle ergometer tests in 2038 men and 898 women over 40 years of age (46.8 ± 6.1 years (mean \pm SD) and 47.5 ± 6.6 years), who volunteered in a sports-medical check-up and who were all active in sports for at least three months in the year preceding the screening (4.3 ± 3.1 hours/week respectively 3.6 ± 2.5 hours/week), were evaluated. The range of maximal values for power output (W_{\max}), heart rate (HR_{\max}), systolic blood pressure (SBP_{\max}) and peak plasma lactate concentrations ($PPLa$) during progressive cycle ergometer testing are presented. W_{\max} varied with sex (male = 1, female = 0), age (year) and height (cm): $W_{\max} = 65.3 \times (\text{sex}) + 2.0 \times (\text{height}) - 1.9 \times (\text{age}) - 67.9$ ($SEE = 38.2$; $r = 0.76$). Furthermore, the weighing of different factors that influence performance was also studied by multiple regression analysis to provide improved precision in standards used to interpret exercise tests. About half of the variation of W_{\max} could be explained by the independent variables age, anthropometric parameters, smoking habits, vital capacity, heart rate, and physical activity parameters. We concluded that being active in endurance sports and/or the use of the bicycle for transport, contributed substantially to cardiovascular fitness in healthy, middle-aged men and women, active in sports.

Introduction

In medical practice progressive maximal exercise tests are often used for diagnostic purposes in patients with cardio-pulmonary complaints as well as for reasons of screening for potential risk factors in healthy persons. Tests for screening or preventive purposes are often performed in persons who are physically active in occupation or leisure time. In this context, physical activities in leisure time are mostly sport activities. In sports-medical practice these tests are also used to judge the aerobic power in relation to training activities.

Measurement of the maximal oxygen consumption ($VO_{2\max}$) is the most commonly preferred index of performance in progressive maximal exercise testing. It is clear that measurements of oxygen consumption are

limited by expense, time and technical assistance and therefore not attractive to use on a large scale. Several investigators introduced an equation to predict VO_2max based on treadmill endurance time (3) or maximal power output (Wmax) achieved during a standard cycle ergometer test (7, 17). Storer et al. (17), who developed prediction equations of VO_2max from Wmax (cycle ergometry), suggested that VO_2max is more closely related to Wmax in cycle ergometry than to treadmill time till exhaustion. Wmax achieved during a progressive maximal cycle ergometer test therefore is an attractive alternative index of the cardiopulmonary functional capacity, when direct VO_2max measurement is not feasible. It has also been shown that the intra-individual variation of Wmax is lower than the intra-individual variation of VO_2max (11).

Reference values of progressive maximal exercise testing in middle-aged, active subjects, especially in women are scarce. Therefore, the purpose of this study was to obtain reference values of maximal aerobic power output during cycle ergometry (Wmax) in men and women older than 40 year and active in sport. The relation between maximal power output and variables such as age, anthropometric parameters, physical activity parameters, smoking habits, vital capacity and heart rate was studied as well.

Subjects and Methods

All subjects volunteered to participate in an experimental sports-medical examination. They were over the age of 40 years and active in sports for at least one hour per week during more than 3 months in the year preceding the screening. The selection of the study population is discussed in Chapter 1. The examinations took place at three Centers for Public Health where similar equipment and protocol were used by three specially trained physicians. A detailed medical and family history was taken in a standardized manner. The physician interviewed the subjects on age (year), alcohol consumption (yes/no), smoking habits (yes/no), medication and physical activity. Physical activity included occupational activity (yes/no), use of the bicycle for transport purposes (yes/no) and sport activities. The number of sports (one, two or more than two), type of sport, frequency (times/week) and duration (hours/week) of participation were recorded. A subject was classified as performing an endurance sport (yes/no) if he/she was active in any sport that uses large muscle groups, that can be maintained for a prolonged period of time and is

rhythmical and aerobic in nature, like long distance running, jogging, cycling, ice-skating, or rowing.

Anthropometric measurements included height (cm), weight (kg) and body fat (%) estimated from the sum of 4 skinfolds (4). Body mass index was calculated from weight and height (kg/m^2). Supine blood pressure was measured from the right upperarm by standard sphygmomanometry and supine heart rate was measured by palpation. The vital capacity (VC) was measured by automatized equipment (Discom, Chest Corporation, Tokyo Japan). To judge if there were any contraindications for the exercise test, information was obtained from medical history, physical examination and a 12-lead resting ECG in each subject. Criteria to exclude subjects from the exercise test were based on recommendations by the American College of Sports Medicine (1).

The exercise test was performed on an electrically braked cycle ergometer (Lode, the Netherlands) with the subject seated upright. The pedal frequency was kept between 60 and 80 rpm. Men started for 5 min at a power output of 100 W and women at 50 W as warming-up. Thereafter, power output was increased by 50 W/2.5 min until a heart rate between 140 and 150 beats/min was reached. Then the increment was 25 W/2.5 min until exhaustion. Maximal power output (W_{max}) was the highest power output that could be maintained for 2.5 min. If the final increment was not fully completed, W_{max} was calculated with a correction for the completed time of that increment. In the third minute of recovery, a blood sample was drawn from a cubital vein for measurement of the plasma lactate concentration (PPLa). Heart rate was determined from a 3-lead (II, V1 and V5) ECG, the highest rate recorded was considered as maximal heart rate (HR_{max}). Tests of subjects who did not meet criteria for maximal effort were excluded. Criteria for maximal effort were a $HR_{\text{max}} \geq 220$ minus age and/or $PPLa \geq 8$ mmol/l for men and a $HR_{\text{max}} \geq 220$ minus age and/or $PPLa \geq 5.5$ mmol/l for women (Chapter 5). Tests of 2038 men and 898 women, not using cardiovascular medication, were analysed. In 1896 men and 826 women venous blood samples for the determination of the peak plasma lactate concentration (PPLa) were obtained. Maximal systolic blood pressure (SBP_{max}) was only recorded if the final work load lasted long enough and accurate measurement was possible. SBP_{max} data were obtained in 742 men and 306 women.

Means and standard deviations of variables were calculated for men and women separately. Differences between groups were tested by Student *t*-tests and Mann-Whitney *u*-tests. Correlation coefficients between W_{max}

and independent variables were calculated and scatterdiagrams were inspected. Simple and multiple regression analyses were performed with Wmax as dependent variable. The decision on how many factors to retain in the final form of the regression model presented in Table 4 was made by using the all-possible subset method. That subset with the largest number of prognostic factors with regression coefficients significantly different from zero ($p < 0.001$) was selected and considered to be the best-fitted model.

Results

Table 1:
Characteristics of the study population (mean \pm SD).

	Men (n=2038)		Women (n=898)
Age (year)	46.8 \pm 6.1	ns	47.5 \pm 6.6
Weight (kg)	77.1 \pm 8.5	*	63.4 \pm 7.6
Height (cm)	175.7 \pm 6.4	*	163.5 \pm 6.0
Body fat (%)	23.5 \pm 5.0	*	33.1 \pm 5.0
Body mass index (kg/m ²)	25.0 \pm 2.2	*	23.7 \pm 2.6
Smokers (%)	26.2	*	18.9
Occupational activity (%)	23.9	*	32.3
Cyclists for transport (%)	46.0	*	65.0
Sport activity:			
number of sports	1.9 \pm 0.8	ns	2.0 \pm 0.8
frequency (times/week)	3.0 \pm 1.9	ns	2.9 \pm 1.9
duration (hours/week)	4.3 \pm 3.1	*	3.6 \pm 2.5
endurance athletes (%)	62.9	*	47.1
vital capacity (l)	5.4 \pm 0.8	*	3.9 \pm 0.6
heart rate (beats/min)	62.6 \pm 9.0	*	67.1 \pm 9.2
blood pressure systolic (mmHg)	134.0 \pm 13.7	*	132.6 \pm 15.1
blood pressure diastolic (mmHg)	88.4 \pm 8.3	*	83.5 \pm 8.2
Maximal exercise:			
Wmax (W)	254.3 \pm 46.1	*	161.7 \pm 28.2
Wmax/kg body weight (W/kg)	3.3 \pm 0.6	*	2.6 \pm 0.5
Wmax FFM ¹⁾ (W/kg)	4.3 \pm 0.8	*	3.8 \pm 0.6
heart rate (beats/min)	175.6 \pm 11.9	*	171.6 \pm 11.7
blood pressure systolic (mmHg)	211.4 \pm 20.5	*	190.0 \pm 19.5
peak plasma lactate (mmol/l)	11.2 \pm 2.8	*	8.4 \pm 2.7
¹⁾ FFM = Fat Free Mass			
* $p < 0.01$, ns = not significant			

Table 2:Mean \pm SD in 5-year age groups in male (M) and female (F).

Age (y)	Sex	Number	Weight (kg)	Height (cm)	Wmax (W)	Wmax/kg (W/kg)	HRmax (b/min)	SBPmax (mmHg)	PPLa (mmol/l)
40-44									
	M	1057	77.5 ± 8.6	176.5 ± 6.5	263.8 ± 45.1	3.4 ± 0.6	178.2 ± 11.7	211.2 ± 19.2	11.5 ± 2.8
	F	426	62.5 ± 7.5	164.5 ± 5.8	171.0 ± 27.1	2.8 ± 0.4	175.2 ± 10.6	186.8 ± 18.1	8.5 ± 2.4
45-49									
	M	530	77.2 ± 8.4	175.6 ± 5.9	253.8 ± 43.7	3.3 ± 0.6	175.8 ± 10.9	210.7 ± 20.7	11.4 ± 2.9
	F	213	63.9 ± 7.8	163.5 ± 5.4	162.7 ± 24.9	2.6 ± 0.5	171.7 ± 10.1	186.5 ± 18.6	8.5 ± 2.6
50-54									
	M	242	76.5 ± 8.3	174.3 ± 6.3	242.5 ± 43.0	3.2 ± 0.6	171.1 ± 11.1	210.6 ± 19.7	10.2 ± 2.6
	F	140	64.2 ± 7.4	163.0 ± 6.1	153.3 ± 25.9	2.4 ± 0.4	167.2 ± 11.9	196.1 ± 19.9	8.2 ± 2.5
55-59									
	M	128	76.3 ± 7.8	174.4 ± 5.9	232.6 ± 40.4	3.1 ± 0.6	169.3 ± 10.2	213.2 ± 22.5	9.9 ± 2.7
	F	76	64.0 ± 7.7	161.4 ± 6.1	143.2 ± 20.8	2.3 ± 0.4	165.6 ± 11.1	195.6 ± 21.6	7.7 ± 2.5
60-64									
	M	57	74.0 ± 8.8	172.2 ± 7.1	206.5 ± 32.0	2.8 ± 0.4	165.6 ± 11.7	210.4 ± 32.6	9.9 ± 2.6
	F	30	67.6 ± 7.1	161.6 ± 7.1	130.3 ± 20.7	1.9 ± 0.3	161.5 ± 10.6	198.5 ± 19.9	6.5 ± 1.4
65+									
	M	24	75.3 ± 7.8	172.0 ± 5.1	193.9 ± 38.3	2.6 ± 0.6	161.7 ± 12.7	228.6 ± 21.5	8.6 ± 3.3
	F	12	61.9 ± 4.9	157.3 ± 7.7	117.6 ± 21.6	1.9 ± 0.4	159.9 ± 7.1	200.8 ± 26.9	6.1 ± 2.2

The characteristics of the study population, shown in Table 1, demonstrate differences between men and women for almost all variables. Women more often said to be physically active at work than men (32% versus 24%), mostly in housekeeping activities. Women also more often used the bicycle for transport than men (65% versus 46%). There was no difference between sexes in the number of sports performed and the frequency of sport activities. Men spent more hours in sport activities than women (4.3 ± 3.1 versus 3.6 ± 2.5 hours/week, $p < 0.001$). Men more often performed endurance sports than women (63% versus 47%). These percentages were 41% and 27% respectively in subjects who were active in only one type of sport.

In Table 2, reference values for a number of variables measured during maximal cycle ergometer testing of the study population are presented in 5-year age groups for men and women. For most variables, except for SBPmax, we found a decrease with increasing age. Gender differences in variables of maximal exercise as presented in the total study population (Table 1) were also present in different age groups (Table 2). Table 2 also shows that the age distribution of the subjects was skewed.

Table 3:

Predictive equations for variables at maximal power output.

Variable	Sex ¹⁾	Height (cm)	Age (year)	Constant	SEE	r
Wmax (W)						
M		2.2	-1.95	-39.5	41.7	0.43
F		1.3	-1.58	+32.7	24.6	0.49
M + F	65.3	2.0	-1.86	-67.9	38.2	0.76
HRmax (beats/min)						
M			-0.60	+203.6	11.3	0.31
F			-0.65	+202.6	10.9	0.37

¹⁾For equation applicable to both sexes male is coded 1, female is coded 0.

Wmax showed significant correlations with sex, age, and size-related variables (Table 3). Weight did not exert a significant effect once age and height had been taken into account in stepwise regression. In Table 3 also the regression equations for HRmax with age as independent variable is given. Of the variables measured, the single variable associa-

ted with the least variance in both sexes was the VC (l): $W_{\max} = 23.3 \times VC + 127.9$ (SEE, 42.2; r , 0.41) in men and $W_{\max} = 18.3 \times VC + 91.0$ (SEE, 25.9; r , 0.39) in women. In Table 4 the results of multiple regression procedures with W_{\max} as dependent variable and eight different independent variables are presented for men and women. For reasons of presentation, the sequence of independent variables in the table does not correspond with the sequence of entrance into the regression models. The regression model for W_{\max} in men is: $W_{\max} (\text{watt}) = 251.5 - 1.78 \times \text{age (years)} + 21.1 \times \text{cycling (yes 1, no 0)} + 23.6 \times \text{endurance sport (yes 1, no 0)} - 15.4 \times \text{smoker (yes 1, no 0)} - 1.27 \times \text{heart rate at rest (beats/min)} + 11.24 \times \text{vital capacity (l)} + 1.32 \times \text{weight (kg)} - 0.76 \times \text{body fat (\%)}$. In men 48% of the variation in W_{\max} could be explained by the independent variables included in the model. In women this was 44%. Variables such as height, body mass index, duration of sport activities (hours/week), occupational activity (yes/no), hemoglobin, alcohol consumption (yes/no) and blood pressure did not contribute significantly in these regression models.

Table 4:
Multiple regression models (regression coefficients (r.c.) and corresponding t-ratios) with W_{\max} as dependent variable in men ($n=2014$) and women ($n=860$).

Variable	W_{\max}			
	Men		Women	
	r.c.	t-ratio	r.c.	t-ratio
Constant	251.49	21.0	193.30	16.8
Age	-1.78	-13.4	-1.41	-11.5
Cycling	21.12	13.9	14.03	9.0
Endurance sport	23.55	14.5	5.18	3.3
Frequency	-	-	1.53	3.7
Smoker	-15.40	-8.9	-	-
Heart rate	-1.27	-14.8	-0.53	-6.5
Vital capacity	11.24	10.0	6.83	4.7
Weight	1.32	11.4	1.06	8.5
Body fat	-0.76	-4.1	-1.23	-6.3
Multiple r	0.69		0.66	
SEE	33.7		21.3	

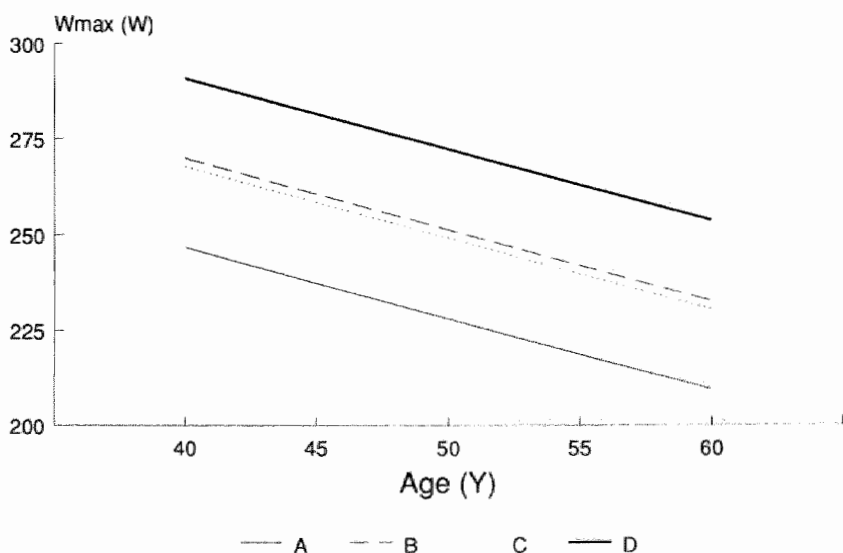
Users of the bicycle for transport (cycling), athletes active in an endurance type of sport (endurance sport) and smokers are coded 0 or 1 depending on a negative resp. positive reaction.

All t-ratios are significant at 0.001 level.

Being active in endurance sports and the use of the bicycle contributed almost equally to W_{\max} in men (Figure 1). In women, the use of the bicycle for transport had a more important influence on W_{\max} than endurance sports (Figure 2). Frequency of sport activities showed better correlations with physical fitness variables than hours active in sport (data not shown). It further appears from Table 4 that the frequency of sport activities made a small but significant contribution to W_{\max} in women, but not in men. Smoking habits were included in the regression model in men, but not in women.

Figure 1:

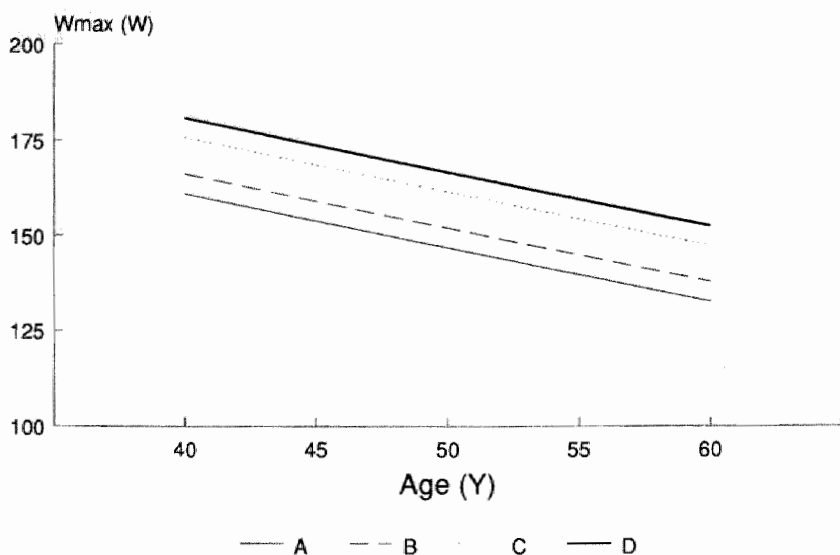
W_{\max} and age in relation with physical activity in men according to the regression analysis in Table 4. All other predicting factors are held constant.



- | | |
|--------|---|
| Line A | use of the bicycle for transport; no (0) |
| | and endurance sport activities; no (0) |
| Line B | use of the bicycle for transport; yes (1) |
| | and endurance sport activities; no (0) |
| Line C | use of the bicycle for transport; no (0) |
| | and endurance sport activities; yes (1) |
| Line D | use of the bicycle for transport; yes (1) |
| | and endurance sport activities; yes (1) |

Figure 2:

Wmax and age in relation with physical activity in women according to the regression analysis in Table 4. All other predicting factors are held constant.



- | | |
|--------|--|
| Line A | use of the bicycle for transport; no (0)
and endurance sport activities; no (0) |
| Line B | use of the bicycle for transport; yes (1)
and endurance sport activities; no (0) |
| Line C | use of the bicycle for transport; no (0)
and endurance sport activities; yes (1) |
| Line D | use of the bicycle for transport; yes (1)
and endurance sport activities; yes (1) |

Discussion

This report presents reference values for maximal power output for progressive maximal aerobic exercise testing on a cycle ergometer in a healthy and active middle-aged population. The importance of cardiovascular fitness is pointed out by several investigators (2, 5) who stated that

a lower level of cardiovascular fitness is associated with a higher risk of death from coronary heart disease and cardiovascular disease in clinically healthy men, independent of conventional coronary risk factors. Cardiovascular fitness in our study population was assessed by measurements of maximal aerobic power output on a cycle ergometer which is a good parameter for aerobic capacity (1, 7, 17).

As a result of promoting physical activity in order to gain higher fitness levels, exercise testing in active, older men and women is becoming more popular and so is the need for reference values. Accepting the idea that reference values should give information about desired levels of a certain tested functional variable, it is reasonable to select a group of physically fit subjects, active in sports, as a reference group in order to make a fair comparison if individuals are tested on this particular function.

Jones et al. (7, 8) and Nordenfelt et al. (14) found lower values of W_{max} in similar age groups compared to results of Weeda (18) and the present study. Differences in selection criteria such as age and the level of physical activity are factors that may explain these findings. Jones et al. (7, 8) studied a selection, although not randomly selected, comparable to the general population with an age range of 15 to 71 years. The study population of Nordenfeldt et al. (14) was a randomly selected group between 20 and 79 years. Therefore, it can be assumed that in these two studies there was no selection bias related to physical activity. Compared with mean values in similar age groups of men in the study of Weeda (18), the men of the present study population reached slightly higher values of W_{max} in the age group of 40 to 44 year and slightly lower values in older age groups. Weeda (18) studied healthy male employees working in the steel industry, with a relatively great percentage of volunteers who were also physically active in leisure time. His cycle ergometer tests were performed with a comparable protocol (30 W/3 min and starting at 90 W for 5 min) until exhaustion according to objective criteria (18). It is obvious that in studies on a relatively active population, and also using criteria for maximal effort, as was done in the present study and in the study of Weeda (18), higher levels of W_{max} are to be expected.

In groups of men of comparable age SBP_{max} values presented by Weeda (18) were 20 to 30 mmHg lower whereas values of W_{max} and HR_{max} were slightly higher than in the present study. The values measured in the present study are somewhat lower (5 to 10 mmHg) than the values of SBP_{max} measured during a near maximal cycle ergometer test (> 90% of the maximal predicted heart rate) in apparently healthy

middle-aged Norwegian men (6). It is likely that technical measurement problems when the subjects become exhausted attributed to differences between values of maximal SBP of the present study and other studies (7, 18). However, a more physiological explanation cannot be excluded.

As can be concluded from above, the maximal power output in a study population is dependent on the selection criteria for this population. In order to collect more information on the influence of different variables on the variation of W_{max} , multiple regression analysis was performed. Age was an important independent factor influencing maximal power output. In the studies of Jones et al. (7, 8) and Weeda (18) the influence of body height on W_{max} was more pronounced than that of body weight, independent of age. Similar regression procedures revealed also better correlations of W_{max} with height than with weight in our group. The regression coefficients as presented in Table 3 were lower than those found by Jones et al. (7), who found a coefficient of 0.86 for W_{max} with sex, age, and height as independent variables, and 0.72 and 0.67 respectively in men and women for W_{max} with age and height as independent variables. These differences in correlations can be explained by differences in distributions of age, height and as mentioned before, physical activity.

In the present study VC was the most powerful predictor of W_{max} in correlation analyses of single variables and the relationship was relatively unaffected by age and sex, which is in agreement with findings of Jones et al. (7). Jones concluded that, since the vital capacity is closely related to other size variables, it may represent a reliable predictor of maximal exercise capacity in subjects free from respiratory disease. The relationship between vital capacity and physical fitness variables may be an explanation for the finding of Kannel et al. (9), in the large prospective Framingham study, that VC was a predictive factor for coronary heart disease. In practice values of the vital capacity are not always available, so other factors such as age and height, which show comparable correlations with W_{max} , are more useful.

The occurrence of other predicting variables in Table 4 are in agreement with findings in the literature on the relationships between single variables and maximal power output. Almost half of the variation in W_{max} can be predicted from these independent variables. The contribution made by these analyses lies in the weighting of different factors that influence performance. Weight became a more important predictor when vital capacity (related to height) was entered in the analysis. Independent of other anthropometric parameters, body fat was an explaining factor of

Wmax. It is difficult to explain single contributions of physiological parameters in the total context of all explaining factors in the models in men and women. It is perhaps of more interest to evaluate differences between male and female subjects, active in sports and the contribution of physical activity parameters in the models. Smoking habit was no explaining factor in women. This may be caused by the fact that not only fewer women were smokers but also that female smokers smoked significantly fewer cigarettes than male smokers (data not shown). The findings in men are in agreement with those of other authors (12, 13, 18), who all reported negative relationships between smoking habits and physical fitness variables (VO_2max or treadmill endurance time). Since the subjects were not allowed to smoke at least three hours before the exercise test, it can be assumed that there was no direct effect of cigarette smoking on the test outcome.

The significant contribution of supine heart rate to the regression equations both in men and women was to be expected. The predicting value of heart rate seems more pronounced in men compared to women. Habitual training enables a person to achieve a certain cardiac output at rest, as well as during exercise, with a lower heart rate and a larger stroke volume. This improves the economy of the heart muscle as far as energy requirement and oxygen demand are concerned (1).

Investigators (16, 18) who classified physical activity in different classes ranging from sedentary to strenuous activities, found an association between physical activity and physical fitness. This association was not expected to be so powerful in our population because only subjects who were active in sports were allowed to participate in the study. Still it appears from Table 4 that in women the frequency of sport activities is an independent explaining factor, indicating that considerable variations exist in sport activities of women in this group. On the other hand the impact of the type of physical activity on maximal power output was found to be more important. It has been shown that elite ball players, and especially those who played as team members, have lower maximal oxygen uptakes than elite endurance athletes (1). Higher levels of aerobic capacity in endurance athletes may be explained by higher initial levels of VO_2max and Wmax by genetic endowment or by a strong hereditary component that influences the functional adaptability (10), resulting in a choice for endurance sports or simply by more aerobic training. In women the contribution of the use of the bicycle for transport to the variation in maximal power output is greater than for example endurance activities in sport. Compared to men, the difference in

contribution of types of physical activities can be explained by more intensive use of the bicycle for transport and fewer intensive endurance activities of women in the study population. This may also explain the additional contribution of the frequency of sport activities in the models in women.

In summary, the reported values of maximal power output, a suitable indicator of aerobic power, of this study population can be considered as normal standards of middle-aged men and women, in sports-medical practice. About half of the variation of maximal power output could be explained by the independent variables age, anthropometric parameters, smoking habits, vital capacity, heart rate and physical activity parameters. In healthy persons active in sport, physical activities not related to sports like use of the bicycle for transport contribute substantially to cardiovascular fitness. This contribution is more pronounced in women than in men.

References

1. Åstrand P-O., Rodahl K.: Textbook of Work Physiology. New York, McGraw-Hill, 1986.
2. Blair S.N., Kohl H.W. III, Paffenbarger R.S. Jr, Clark D.G., Cooper K.H., Gibbons L.W.: Physical fitness and all cause mortality: a prospective study of healthy men and women. *JAMA* 262: 2395-2401, 1989.
3. Bruce R.A., Kusumi F., Hosmer D.: Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *Am Heart J* 85: 546-562, 1973.
4. Durnin J.V.G.A., Womersley J.: Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged 16 to 72 years. *Br J Nutr* 32: 77-97, 1974.
5. Ekelund L-G., Haskell W.L., Johnson J.L., Whaley F.S., Criqui M.H., Sheps D.S.: Physical fitness as a predictor of cardiovascular mortality in asymptomatic north american men. *N Engl J Med* 319: 1379-1384, 1988.
6. Erikssen J., Jervell J., Forfang K.: Blood pressure responses to bicycle exercise testing in apparently healthy middle-aged men. *Cardiology* 66: 56-63, 1980.
7. Jones N.L., Makrides L., Hitchcock C., Chypchar T., McCartney N.: Normal standards for an incremental progressive cycle ergometer test. *Am Rev Respir*

Dis 131: 700-708, 1985.

8. Jones N.L., Summers E., Killian K.J.: Influence of age and stature on exercise capacity during incremental cycle ergometry in men and women. *Am Rev Respir Dis* 140: 1373-1380, 1989.
9. Kannel W.B., Hubert H., Lew E.A.: Vital capacity as a predictor of cardiovascular disease: The Framingham study. *Am Heart J* 102: 311-315, 1983.
10. Klissouras V.: Genetic limit of functional Adaptability. *Int Z Angew Physiol* 30: 85-94, 1972.
11. Kuipers H., Verstappen F.T.J., Keizer H.A., Geurten P., Kranenburg G.: Variability of aerobic performances in the laboratory and its physiological correlates. *Int J Sports Med* 6: 197-201, 1985.
12. McDonough J.R., Kusumi F., Bruce R.A.: Variations in maximal oxygen intake with physical activity in middle-aged men. *Circulation* 61: 743-751, 1970.
13. Milesis C.A.: Prediction of treadmill performance from clinical characteristics in healthy persons. *J Cardiopulm Rehab* 7: 365-373, 1987.
14. Nordenfeldt I., Adolfsson L., Nilsson J.E., Olsson S.: Reference values for exercise tests with continuous increase in load. *Clin Physiol* 5: 161-172, 1985.
15. Patton J.F., Vogel J.A., Mello R.P.: Evaluation of a maximal predictive cycle ergometer test of aerobic power. *Eur J Appl Physiol Occup Physiol* 49: 131-140, 1982.
16. Sallis J.F., Patterson T.L., Buono M.J., Nader P.R.: Relation of cardiovascular fitness and activity to cardiovascular disease risk factors in children and adults. *Am J Epidemiol* 127: 933-941, 1988.
17. Storer T.W., Davis J.A., Caiozzo V.J.: Accurate prediction of $\text{VO}_{2\text{max}}$ in cycle ergometry. *Med Sci Sports Exerc* 22: 704-712, 1990.
18. Weeda H.W.H.: Maximal exercise tests in 331 male employees; relations to age and physical activity: Dissertation. Haarlem, De Vrieseborch, 1989.
19. Whipp B.J., Koyal S.N., Wasserman K.: Anaerobic threshold and O_2 uptake kinetics for work increments of various durations. *Med Sci Sports Exerc* 6: 67-68, 1974.

40+

Chapter 7

The inter-relationship between cardiovascular fitness, physical activity and cardiovascular risk factors in middle-aged men and women, active in sports

by

A.M.P.M. Bovens, M.A. van Baak,
J.G.P.M. Vrencken, J.A.G. Wijnen,
W.H.M. Saris and F.T.J. Verstappen

Abstract

The association of several indicators of physical activity and cardiovascular fitness with risk factors such as blood pressure, body composition and smoking habits was evaluated in a middle-aged population, active in sports. Data records of 2009 men and 898 women, all volunteers over 40 years (46.8 ± 6.1 years (mean \pm SD) and 47.5 ± 6.6 years respectively) and active in sports (4.3 ± 3.1 hours/week and 3.6 ± 1.9 hours/week respectively), were available. Physical activity was recorded as sport activity (number and type of sport, frequency and duration), occupational activity and the use of the bicycle for transport. Cardiovascular fitness was expressed as maximal power output (W_{\max}/kg) achieved during a progressive maximal cycle ergometer test. Pearsons product moment correlations between cardiovascular fitness and risk factors indicated significantly lower risks among more fit individuals. Except for smoking habits all risk factors showed better associations with fitness in women than in men. The association of physical activity indicators and risk factors was less strong than the fitness-risk factor association in both genders. In both genders, joggers had the most favourable risk profile compared to subjects who were active in one of the other four most popular sports. We conclude that in this relatively active and healthy population comparable associations of physical activity and cardiovascular fitness with risk factors for cardiovascular disease were found as in studies on less active populations; physically more fit and physically more active subjects have more favourable cardiovascular risk profiles.

Introduction

In the past a number of studies have shown an inverse relationship between coronary heart disease and habitual physical activity (11, 12). Morris et al. (11) concluded that activities other than vigorous aerobic exercises were not inversely associated with risk for coronary heart disease. The inverse relation between physical activity and risk factors for cardiovascular disease is well-established (7, 9, 14). In a review by Powell et al. (13), with a critical analysis of all available studies about physical inactivity as risk factor, it was concluded that physical inactivity can be considered as a risk factor with nearly the same odds ratio as the known risk factors such as hypercholesterolaemia and smoking. However, due to the fact that a larger portion of the general population is exposed

to physical inactivity, the impact on coronary heart disease is considerably larger than that of most of the other known risk factors. Physical activity is an important determinant of physical fitness, so to some extent, fitness is an objective marker for habitual physical activity (2). The measurement of physical activity is a tedious and more difficult task (15) than that of physical fitness, and therefore it is not surprising that a stronger association is found between fitness and cardiovascular risk factors than between physical activity and these risk factors (12). There is also evidence that fitness in itself is inversely related to cardiovascular disease and all-cause mortality independent of other risk factors (3, 6). It is clear that these investigations contributed to the acceptance of regular exercise as a way to promote health and to reduce the risk of cardiovascular disease and other chronic diseases. On the other hand there is an increasing demand for studies on the amount and intensity of physical activity that is needed to achieve positive health effects. The purpose of this study was to evaluate the inter-relationship between several indicators of physical activity, including the type of sport activity and cardiovascular fitness, and cardiovascular risk factors such as blood pressure, body composition, and smoking habits, in a physically active population. The study population consisted of volunteers over 40 years of age, active in sports.

Methods

At three different Centers of Public Health in the south of the Netherlands 3918 middle-aged volunteers participated in an experimental sports-medical examination. They were all active in sports for at least one hour/week over a period of three months in the year preceding the examination. A minority (30%) was active in competitive sports.

At the three locations the same protocol and equipment were used by three specially trained physicians throughout the study period of 3.5 years. A detailed medical and family history was taken by interview. The subjects were asked about their age, smoking habits, and medication. Detailed information was asked about their physical activity. Physical activity was recorded as sport activity, occupational activity (active/inactive, reported by the subjects), and regular use of the bicycle for transport (yes/no). Number of sports were recorded, with a maximum of three sports on which further information was collected on duration and frequency per week in the year preceding the interview. Duration and

frequency were averaged per year.

Anthropometric measurements taken included body height, body weight and body fat (%), estimated from the sum of four skinfolds (5). Furthermore the body mass index ($\text{BMI} = \text{weight (kg)}/\text{height (m)}^2$) was used as a second measure of body fatness. Supine blood pressure was measured from the right upper arm by standard sphygmomanometry. If it was elevated, it was remeasured after at least another 5 min supine rest. The lowest value was taken. Fifth-phase diastolic pressure was used.

If no contra-indications were present (1), a maximal cycle ergometer test was performed to assess cardiovascular fitness. The subject was seated upright on an electrically braked cycle ergometer (Lode, the Netherlands) with 3-lead (II, V1, and V5) ECG electrodes in place. The test was performed with a pedal frequency between 60 and 80 rpm, starting 5 min at a power output of 100 W for men and 50 W for women. Thereafter, the power output was increased by 50 W/2.5 min until a heart rate was reached between 140 and 150 beats/min. Then the increment was 25 W/2.5 min until exhaustion. Maximal power output (W_{max}) was the highest power output that could be maintained for 2.5 min. If the final power output was not completed, W_{max} was calculated out of the completed time. Heart rate was determined from the ECG at the end of each power output. Maximal heart rate was the highest heart rate recorded. In the third min of recovery a blood sample was drawn from a cubital vein for measurement of the plasma lactate concentration (LA 640 Roche) (8). Only subjects who fulfilled criteria for a maximal exercise performance were included in the analysis. These criteria were a maximal heart rate ≥ 220 minus age and/or peak plasma lactate concentration ≥ 8 mmol/l in men and ≥ 5.5 mmol/l in women (Chapter 5).

Complete data records of 2009 men and 898 women were available. Subjects using cardiovascular medication were excluded from the analysis. Results are presented as means \pm standard deviations. For statistical analysis Pearson correlation and partial correlation procedures were used. Differences between subpopulations were tested by analysis of covariance, two-sample t-tests and Mann-Whitney u-tests. $p < 0.05$ was regarded as statistically significant.

Results

Men reported significantly more sport activities in hours/week than women (Table 1). Frequency and number of sport activities did not

differ between genders. Men had higher values of systolic and diastolic blood pressure and more men reported to be a current smoker (26% versus 19%).

Table 1:
Characteristics (mean \pm SD) of study population.

	Men (n=2038)	Women (n=898)	
Age (year)	46.8 \pm 6.1	47.5 \pm 6.6	**
Body fat (%)	23.5 \pm 5.0	33.1 \pm 5.0	***
BMI (kg/m ²)	25.0 \pm 2.2	23.7 \pm 2.6	***
SBP (mmHg)	135.7 \pm 14.0	132.7 \pm 15.0	***
DBP (mmHg)	88.4 \pm 8.3	83.5 \pm 8.2	***
Smoking (%)	26.2	18.9	***
Wmax (W)	254 \pm 46	162 \pm 28	***
Wmax/kg (W/kg)	3.3 \pm 0.6	2.6 \pm 0.5	***
Wmax/kg FFM (W/kg) ¹⁾	4.3 \pm 0.8	3.8 \pm 0.6	***
Number of sports	1.9 \pm 0.8	2.0 \pm 0.8	ns
Duration (hours/week)	4.3 \pm 3.1	3.6 \pm 2.5	***
Frequency (times/week)	3.0 \pm 1.9	2.9 \pm 1.9	ns
Occupational activity (%)	23.9	32.3	***
Bicycle for transport (%)	46.0	65.0	***

¹⁾FFM = Fat Free Mass
 p<0.01 (Mann-Whitney u-test), *p<0.001 (t-test),
 ns = not significant

In tables 2 and 3 Pearsons product moment correlations are presented for men and women separately. All correlations between cardiovascular fitness expressed as Wmax/kg and risk factors indicated significantly lower levels of risk among more fit individuals. Since age and body mass index were highly correlated with Wmax/kg, the fitness-risk factor association was analysed adjusting for age and body mass index. The relations between fitness and body fat as well as blood pressure became less strong but remained significant, whereas the correlation coefficients with smoking habits in men as well as in women increased. Except for smoking habits all risk factors showed better associations with Wmax/kg in women than in men.

Table 2:

Correlations between cardiovascular fitness (Wmax/kg), physical activity, and cardiovascular risk factors in 2009 men active in sport.

Variable	Fitness		Physical activity		
	Wmax/kg	Wmax/kg (adj.) ¹⁾	Sport times/wk	Occup. activity	Bicycle transport
Body fat (%)	-0.35 ***	-0.12 ***	-0.10 ***	-0.11 ***	0.01 ns
SBP (mmHg)	-0.17 ***	-0.06 **	0.02 ns	-0.02 ns	0.02 ns
DBP (mmHg)	-0.18 ***	-0.06 **	-0.03 ns	0.01 ns	-0.02 ns
Smoking (%)	-0.19 ***	-0.25 ***	-0.12 ***	0.02 ns	-0.04 *
BMI (kg/m ²)	-0.40 ***		-0.12 ***	0.03 ns	-0.06 **
Wmax/kg adj. ¹⁾ (W/kg)			0.23 ***	0.01 ns	0.26 ***

ns = not significant, *p<0.05, **p<0.01, ***p<0.001
¹⁾adj. = adjusted for age and BMI

Table 3:

Correlations between cardiovascular fitness (Wmax/kg), physical activity, and cardiovascular risk factors in 898 women active in sport.

Variable	Fitness		Physical activity		
	Wmax/kg	Wmax/kg (adj.) ¹⁾	Sport times/wk	Occup. activity	Bicycle transport
Body fat (%)	-0.51 ***	-0.12 ***	-0.10 ***	-0.18 ***	0.10 ***
SBP (mmHg)	-0.30 ***	-0.10 **	0.01 ns	-0.05 ns	-0.03 ns
DBP (mmHg)	-0.26 ***	-0.08 *	-0.02 ns	-0.04 ns	-0.01 ns
Smoking (%)	-0.06 *	-0.13 ***	-0.07 *	0.04 ns	-0.11 ***
BMI (kg/m ²)	-0.50 ***		-0.08 *	-0.02 ns	0.02 ns
Wmax/kg adj. ¹⁾ (W/kg)			0.20 ***	-0.05 ns	0.21 ***

ns = not significant, *p<0.05, **p<0.01, ***p<0.001
¹⁾adj. = adjusted for age and BMI

The relation between physical activity indicators and risk factors was less powerful than the fitness-risk factor association in both genders. Body fat and smoking habits turned out to be significantly related with sport activity indicators in men and women. Occupational activity showed a negative association with body fat in women and to a lesser degree in men (Tables 2 and 3).

The association between physical activity in leisure time and cardiovascular fitness was moderately strong in men and women, whereas no significant association between occupational activity and cardiovascular fitness was found in either gender.

Table 4:

Characteristics of subpopulations of men, active in at least three sports (3) or active in only one type of sport and one of the 5 most popular sports (A-E).

	Active in at least 3 sports n=504 3	Active in just one type of sport				
		cycling n=152 A	jogging n=133 B	tennis n=106 C	soccer n=95 D	volleyball n=48 E
Age (year)	47.4 ±6.7	47.9 ±7.5	45.1 ±4.7	47.2 ±5.5	44.8 ±4.8	45.7 ±5.2
Body fat (%)	23.5 ±5.0	23.1 ±5.1	21.1 ±4.9	24.8 ±4.8	23.4 ±4.8	23.8 ±4.9
BMI (kg/m ²)	24.7 ±2.2	25.1 ±2.3	24.3 ±2.1	25.2 ±2.1	25.9 ±2.0	24.4 ±2.1
SBP (mmHg)	136.0 ±13.0	135.2 ±12.2	133.8 ±11.6	135.8 ±13.9	135.1 ±13.3	135.1 ±10.9
DBP (mmHg)	86.1 ±7.5	86.6 ±7.2	85.2 ±7.2	86.1 ±7.8	86.5 ±7.7	86.8 ±7.4
Smoker (%)	22.0	27.0	19.5	39.6	36.8	29.2
Wmax (W)	260.8 ±48.0	269.9 ±53.1	270.3 ±35.5	231.3 ±37.7	244.3 ±36.5	240.0 ±39.6
Wmax/kg (W/kg)	3.4 ±0.6	3.5 ±0.7	3.6 ±0.5	3.0 ±0.5	3.1 ±0.5	3.2 ±0.4
Duration (hours/week)	6.1 ±3.4	4.8 ±3.7	3.6 ±2.4	3.0 ±3.0	1.7 ±1.4	1.7 ±1.1
Frequency (times/week)	4.3 ±2.2	2.3 ±1.5	3.2 ±1.5	1.9 ±1.0	1.2 ±0.8	1.1 ±0.4
Occupational activity (%)	21.6	25.3	22.0	12.5	44.1	19.6
Bicycle for transport (%)	48.4	49.3	47.4	33.0	51.6	56.3

Table 5:

Characteristics of subpopulations of women, active in at least three sports (3) or active in only one type of sport and one of the 5 most popular sports (A-E).

	Active in at least 3 sports	Active in just one type of sport				
	n=252 3	cycling n=70 A	jogging n=55 B	tennis n=45 C	soccer n=17 D	volleyball n=12 E
Age	47.6	46.6	50.3	43.6	47.5	50.8
(year)	±6.6	±5.2	±8.7	±3.7	±5.3	±8.6
Body fat	33.2	32.3	34.6	32.1	35.5	34.3
(%)	±4.9	±4.1	±4.3	±4.1	±5.3	±5.2
BMI	23.6	23.4	24.2	23.3	25.3	23.5
(kg/m ²)	±2.8	±2.3	±2.8	±2.3	±3.0	±1.8
SBP	131.8	132.3	136.2	127.9	133.8	138.3
(mmHg)	±14.7	±12.7	±14.8	±14.0	±13.9	±16.3
DBP	82.8	84.1	85.5	82.4	87.1	82.9
(mmHg)	±7.6	±9.4	±8.1	±7.6	±6.6	±6.2
Smoker (%)	17.1	27.1	25.5	22.2	23.5	8.3
Wmax	167.5	154.7	143.1	177.2	158.6	145.0
(W)	±27.5	±25.2	±30.3	±24.4	±32.7	±21.1
Wmax/kg	2.7	2.5	2.3	2.9	2.4	2.4
(W/kg)	±0.5	±0.4	±0.5	±0.4	±0.5	±0.4
Duration	4.9	3.0	1.0	2.7	1.5	3.4
(hours/week)	±2.5	±1.8	±0.6	±1.4	±0.9	±2.5
Frequency	4.1	2.0	1.0	2.7	1.7	2.2
(times/week)	±2.0	±1.9	±0.6	±1.1	±1.3	±2.1
Occupational activity (%)	25.0	47.1	40.0	24.4	25.0	16.7
Bicycle for transport (%)	70.6	54.3	67.3	62.2	58.8	58.3

The effect of the type of sport on the risk factors, cardiovascular fitness and physical activity was studied in subpopulations of athletes who were active in only one sport. A selection was made of the 5 most popular sports in each gender (Table 4 and 5). In both genders also subjects, who were active in at least 3 different types of sport, were studied. Male

subjects, active in endurance types of sport (joggers and cyclists) had the highest values for maximal power output and sport activity (analysis of covariance and two-sample t-tests). Joggers attained the highest maximal power output in women. In both genders joggers had the most favourable profiles of risk factors compared to subjects active in one of the five other sports and even compared to subjects who were active in three sports or more. Analysis of covariance between subpopulations of subjects, active in one type of sport revealed significant differences for body fatness in both genders and for smoking habits in men (data not shown). Blood pressure did not differ between types of sport in either genders. In general the specific sport groups differed significantly from each other with respect to aerobic power and frequency and duration of sport activities.

Discussion

The results indicated that even in a fit and active population physically more fit and physically more active subjects have better cardiovascular risk profiles based on body fat, blood pressure and smoking habits. Furthermore, it was found that cardiovascular fitness, expressed as W_{max}/kg , correlates better with cardiovascular disease risk factors such as blood pressure, body composition characteristics and smoking habits than physical activity indicators with these factors. This is in agreement with findings in the study of Sallis et al. (14). The comparative strength of fitness-risk factor and activity-risk factor relations has seldom been studied. Caspersen (4) enumerated several methodological problems in physical activity epidemiology. Because physical activity is a complex behavior, problems arise when making operational definitions leading to a wide variation in physical activity indicators (4). Furthermore, it was stated by Saris (15) that measurement of physical activity is one of the most difficult tasks of research in exercise science. So it is not surprising that better correlations are found with cardiovascular fitness, mostly as aerobic power, which can be measured more objectively. On the other hand it may be possible that the relationships alter when physical activity parameters can be recorded more precisely or when cardiovascular fitness is difficult to assess. For example, one can imagine that the assessment of cardiovascular fitness in the elderly can be difficult and that in these age groups better correlations with physical activity parameters can be found.

The strength of relations found between cardiovascular fitness and cardiovascular risk factors is comparable to that found in other studies (9, 14). In these studies the cardiovascular fitness indicators (treadmill time (9) and predicted maximal oxygen consumption (14)) were adjusted for the influence of age and body mass index because both variables highly correlated with the fitness indicators. After adjusting for body mass index and age in our study, significant correlations remained between W_{max}/kg and body fat in both genders, indicating that body fat is an independent anthropometric parameter inversely related to fitness. The fitness-blood pressure correlations were reduced after adjusting for body mass index and age, as found by Sallis et al. (14). These results suggest that at least part of the relation can be explained by differences in body composition. Sallis et al. (14) also reported better correlations between fitness and systolic and diastolic blood pressure in women compared to men, which is in agreement with our findings. It is not clear whether this difference is due to differences in body composition between genders. Perhaps the change in blood pressure to training is different in women than in men, which has been suggested by Hagberg (10). He stated that endurance exercise training may elicit greater blood pressure reductions in females than males.

The negative association between smoking habits, recorded as current smokers, and cardiovascular fitness was more pronounced in men than in women. The demonstrated association between smoking habits and fitness is in agreement with findings in the literature (9). Since it was not allowed to smoke 3 hours before the exercise test, no direct effect of cigarette smoking on maximal power output was to be expected. Male joggers have the highest fitness levels and the lowest percentage of current smokers compared to men active in other types of sport. It is not possible to explain these findings on basis of the collected data. Perhaps, more active people may be encouraged to stop smoking as a result of starting and maintaining exercise programs.

Different categories of physical activity may relate to specific aspects of health and therefore it may be important to consider them separately in epidemiologic research (4). In the present study physical activity was defined as sport activity (duration and frequency), occupational activity, and the use of the bicycle for transport. Besides the already mentioned methodological problems in the assessment of physical activity, Folsom et al. (7) mentioned the influence of the small range in physical activity in a population-based sample. Only 34% of men and 17% of women expended 2000 kcal or more per week in leisure time physical activity

(7). A relatively small range of physical activity was expected in the present study population due to the fact that the subjects had to be active in sports for at least three months in the year preceding of the screening. Still a great variance in duration, frequency and intensity of the sport activities was found. For example, the duration of sport activities ranged from 0.25 hours/week to more than 20 hours/week and the intensity of sport activities varied from bowling to competitive triathlon. So even in a population with a considerable variation in physical activity comparable low correlations were found between physical activity and cardiovascular risk factors compared to a population with a smaller range of activity levels (7).

No significant correlation between blood pressure and any index of physical activity was found. This finding was in agreement with other studies (7, 12). In addition to the negative association between fitness and smoking habits also a negative association was found between smoking habits and sport activity (frequency and duration). Both associations were stronger in men compared to women. Surprisingly, the association between smoking habits and the use of the bicycle for transport was stronger in women. The use of the bicycle for transport was found to be an independent predictor of maximal power output (Chapter 6). The contribution was more pronounced in women compared to men suggesting that, apart from sports activities, a more healthy 'way of life' may explain the difference in contribution between genders. This is perhaps also reflected in a more pronounced negative association between smoking habits and the use of the bicycle in women.

When the type of sport is taken into account it appears from Tables 4 and 5 that of the five most popular sports in male as well as in female joggers are characterized by the most favourable values of body composition and the highest mean values of maximal power output and duration and frequency of sport activity. It is not clear whether this is due to a causal relationship between type of sport and the risk factors, because of the cross-sectional nature of the study. Maybe subjects with a more favourable body composition are more likely to participate in running activities, whereas (female) swimmers, characterized by the highest values of body fat and body mass index, are active in that type of sport because in water they are less bothered by their body weight. The positive relation between body fat and the use of the bicycle for transport in females may be explained by the same mechanism. On the other hand it is well known that the endurance type of exercise leads to more improvement of aerobic power (1) and reduction of blood pressure (10)

than more static muscle activities as weight-lifting. This may also explain the results of the subjects who were active in at least three different types of sport, with no better profiles compared to subjects just active in jogging. Furthermore, subjects who are active in more sports are perhaps more interested in social activities than joggers, training mostly individually. We can only speculate on possible explanations because the group of subjects active in more sports is very heterogeneous.

No association was found between cardiovascular fitness and occupational activity in both genders. This has previously been shown by Sobolsky et al. (13) in mainly sedentary but healthy middle-aged workers. The industrial mechanization reduced energy expenditure at work. Obviously more effect can be expected of leisure time physical activities. Furthermore, no objective measures of occupational activities were collected making this variable less valid.

From the results in this study it can be concluded that in an active population, the strength of the association between cardiovascular fitness, physical activity and cardiovascular risk factors is similar to that in studies on less active men and women; physically more fit and active subjects have better cardiovascular risk profiles. Of all sports participants, male and female joggers showed the most favourable risk profiles.

References

1. American College of Sports Medicine: Guidelines for Exercise Testing and Prescription. Philadelphia, Pa: Lea & Febiger; 1986.
2. Blair S.N.: Physical activity leads to fitness and pays off. *Physician Sports Med* 13: 145-150, 1985.
3. Blair S.N., Kohl H.W. III, Paffenbarger R.S., Clark D.G., Cooper K.H., Gibbons L.W.: Physical fitness and all-cause mortality; A prospective study of healthy men and women. *JAMA* 262: 2395-2401, 1989.
4. Caspersen C. J.: Physical Activity Epidemiology; Concepts, Methods, and Applications to Exercise Science. *Exercise and Sports Reviews* 17: 423-473, 1989.
5. Durnin J.V., Womersley J.: Body fat assessed from total body density and its estimation from skinfold thickness. *Br J Nutr* 32: 77-97, 1974.
6. Ekelund L-G, Haskell W.L., Johnson J.L., Whaley F.S., Criqui M.H., Sheps D.S.: Physical fitness as a predictor of cardiovascular mortality in asymptomatic north american men. *N Engl J Med* 319: 1379-1384, 1988.

7. Folsom A.R., Caspersen C.J., Taylor H.L., Jacobs D.R. Jr, Luepker R.V., Gomez-Marín O., Gillum R.F., Blackburn H.: Leisure time physical activity and its relationship to coronary risk factors in a population-based sample. *Am J Epidemiol* 121: 570-579, 1985.
8. Geyssaut A.D., Dormois D., Barthelemy J.C., Lacour J.R.: Lactate determination with the lactate analyser LA 640: a critical study. *Scand J Clin Invest* 45: 145-149, 1985.
9. Gibbons L.W., Blair S.N., Cooper K.H., Smith M.: Association between coronary heart disease risk factor and physical fitness in healthy adult women. *Circulation* 67: 977-983, 1983.
10. Hagberg J.M.: Exercise, Fitness, and Hypertension. In Bouchard C., Shephard R.J. (eds): *Exercise, Fitness, and Health*. Champaign IL, Human Kinetics Books: 455-462, 1988.
11. Morris J.N., Clayton D.G., Everitt M.G., Semmence A.M., Burgess E.H.: Exercise in leisure time: coronary attack and death rates. *Br Heart J* 63: 325-334, 1990.
12. Paffenbarger R.S. Jr, Hyde R.T., Wing A.L., Steinmetz C.H.: A natural history of athleticism and cardiovascular health. *JAMA* 252: 491-495, 1984.
13. Powell K.E., Thompson P.D., Caspersen C.J., Kendrick J.S.: Physical activity and the incidence of coronary heart disease. *Ann Rev Publ Health* 8: 253-287, 1987.
14. Sallis J.F., Patterson T.L., Buono M.J., Nader P.R.: Relation of cardiovascular fitness and physical activity to cardiovascular disease risk factors in children and adults. *Am J Epidemiol* 127: 933-941, 1988.
15. Saris W.H.: The assessment and evaluation of daily physical activity in children: a review. *Acta Paediatr Scand Suppl* 318: 37-48, 1985.
16. Sobolsky J.C., Kolesar J.J., Kornitzer M.D., De Backer G.G., Mikes Z., Dramaix M.M., Degre S.G., Denolin H.F.: Physical fitness does not reflect physical activity patterns in middle-aged workers. *Med Sci Sports Exerc* 20: 6-13, 1988.

40⁺

Chapter 8

General discussion and conclusions

The aim of the present dissertation was to evaluate measurements of a preventive sports-medical screening. This examination was performed in about 4000 men and women as part of an extensive epidemiological study in order to support governmental departments in their policy on preventive (sports-)medicine. The selection mechanism that may have implications for generalization are discussed. The examinations were performed by three physicians at three Centers for Public Health with implications for variability and reliability of measurements. It appeared that the study population was relatively active and healthy, and therefore comparisons with values of reference populations revealed favourable values for most of the subjects. Furthermore it was interesting to investigate the inter-relationships between cardiovascular fitness, physical activity and some cardiovascular disease risk factors in active and healthy middle-aged men and women.

Chapter 1 can be regarded as a general introduction to the other chapters and therefore, the variables which are relevant for the interpretation of results discussed in the other chapters are described. The study population consisted of 3918 volunteers (2661 men and 1257 women). The men were on average 47.3 ± 6.7 years old, the women were 48.1 ± 7.2 years. They were respectively 4.3 ± 3.6 and 3.5 ± 2.8 hours/week active in sports. They underwent an extensive sports-medical check-up including a medical and physical activity history, physical examination, laboratory tests with a maximal exercise test on a cycle ergometer and an evaluation of the test results along with health and sport related advices. Compared to a representative sample of persons, also active in sports, of the same age group, and living in the same region, the study population was younger and consisted of more males. They spent more hours per week on sport activities and were more often active in sports with higher intensities. Perhaps because of this, higher incidence rates of sport injuries were reported compared to a representative sample. The consequences of the selective recruitment of subjects have to be taken into account for the interpretation of the results discussed in this dissertation and the epidemiological study.

In Chapter 2 and Chapter 3 the variability and reliability of flexibility measurements, pulmonary function tests, and maximal exercise testing is studied. The purpose of these studies was to evaluate whether the results collected at different locations with different equipment and observers could be used in the cross-sectional study. The measurements for these studies were taken on five groups of 9-10 volunteers (48 ± 7 years) who were examined at all three laboratories. Each subject (43 men and 5 women) was tested three times within three weeks, the five groups were tested over a period of 2.5 years.

Although simple joint measurements with or without the use of a goniometer are widely used in a sports medicine practice, little is known about the variability and reliability of these measurements. Therefore, along with the measurements of the five groups of middle-aged subjects, nine flexibility measurements were also carried out on eight healthy subjects (5 men and 3 women), aged 30 ± 6 years, to investigate the intratester as well as the intertester variability and reliability of the three observers. The subjects, all volunteers belonging to the university personnel, underwent standardized measurements in eight sessions over a period of 3 months. The intertester variability and reliability were also determined in the five groups as described above. The results indicate that it is difficult to demonstrate either an improvement or worsening of a joint motion of less than 5° to 10° for most joints measured by the same observer (Chapter 2). Furthermore, the intertester variation was not consistent over a longer period of time, so differences between observers cannot be corrected on the basis of a single study at a single point in time. It was also shown that the reliability of most goniometric measurements increases when the observers have more experience.

In Chapter 3 the variability and reliability of three pulmonary function tests (vital capacity, expiratory volume in the first second and the peak flow rate), of measurements during a maximal exercise test on a cycle ergometer (maximal work output, maximal systolic blood pressure, and maximal heart rate), of blood pressure at rest, and percentage of body fat, estimated from the sum of four skinfolds, are evaluated. The variation of these physiological parameters for a given person is the result of several sources: the intertester variation, technical error, and the biological variation of the subject. The results indicated that data on the studied variables could be used for a cross-sectional population study, because the variation between different locations was relatively small compared to the variation between the subjects. In maximal exercise testing the biological variation seems to dominate in the total variation

between laboratories.

Pulmonary function was studied in Chapter 4. The aim of the study presented in this chapter was to investigate whether an active, healthy, and middle-aged population had better pulmonary function than predicted from reference values, which has been demonstrated previously in younger athletes. The vital capacity (VC), the forced vital capacity (FVC), the forced expiratory volume in 1 sec (FEV1.0), the FEV1.0/VC ratio and the peak expiratory flow (PEFR) were determined in a subpopulation, free of signs or symptoms of chronic obstructive lung disease. All measurements were performed by identically instructed and trained personnel with similar electronic spirometry systems. Higher values of VC, FVC, FEV1.0 and PEFR were found compared to generally accepted reference values. The FEV1.0/VC ratio did not show a difference in either gender. Male and female subjects active in endurance types of sport activities had higher volumes (VC, FVC and FEV1.0) than non-endurance athletes. Based on these findings, new regression equations were computed for healthy, middle-aged men and women, active in sports according to smoking habits. Values based on these equations can be regarded as optimal values of pulmonary function in a healthy population.

An important part of the sports-medical check-up is the maximal exercise test, to evaluate the response of the cardiopulmonary system to exercise. Furthermore, it is a valid method to establish the cardiovascular fitness. Cardiovascular fitness was in this dissertation defined as maximal work output (W_{max}) achieved during a graded maximal exercise test on a cycle ergometer. The cycle ergometer test till exhaustion was performed in according to the following protocol: increments of 50 W/ 2.5 min, starting at 100 W in men and 50 W in women, and 25 W/ 2.5 min after a heart rate between 140 and 150 beats/min was reached. W_{max} was the highest work output that could be maintained for 2.5 min or W_{max} was calculated with a correction of the completed time if the final work output was not fully completed. The highest value of the heart rate during the test was considered to be HR_{max} . In the third minute of recovery, a blood sample was drawn from a cubital vein for measurement of the peak plasma lactate concentration (PPLa). It is generally accepted that maximal effort is confirmed by levelling-off of the oxygen consumption with increasing power output. If no gas analysis equipment is available, HR_{max} and PPLa are often used. In Chapter 5 these criteria for maximal effort are discussed with special attention to the difference in PPLa between genders. A significant difference in PPLa of about 2.5

mmol/l, independently from age and maximal heart rate, was found. Possible explanations for this finding were discussed in this chapter. This finding was used in all other analyses of this dissertation on maximal exercise testing in combination with HRmax, as criteria for maximal effort.

In Chapter 6 the range of maximal values for power output, heart rate, systolic blood pressure and peak plasma lactate concentration of the study population, collected during the maximal cycle ergometer test, are presented. Another purpose of the study presented in this chapter was to investigate the relation between Wmax and independent variables such as age, anthropometric parameters, physical activity parameters, smoking habits, vital capacity and heart rate. Multiple regression analysis showed that about half of the variation of the maximal power output could be explained by these independent variables. Furthermore, it appeared that in healthy persons, active in sports, physical activities not related to sports, like the use of the bicycle for transport, contribute substantially to cardiovascular fitness. This contribution was more pronounced in women.

In Chapter 7, the inter-relationship of cardiovascular fitness, several indicators of physical activity and cardiovascular risk factors such as blood pressure, body composition and smoking habits were evaluated. Physical activity was recorded as sport activity (number and type of sport, frequency and duration), occupational activity (active or inactive) and the use of the bicycle for transport. Correlations between cardiovascular fitness and risk factors indicated lower risks among more fit individuals. Except for smoking habits all risk factors showed better associations with fitness in women than in men. In both genders, the association of physical activity indicators and risk factors was less strong than the fitness-risk factor association. The finding that in both genders, joggers had the most favourable risk profiles compared to subjects who were active in one of the other four most popular sports, was also discussed in this chapter. It was concluded that in a relatively active and healthy population similar associations of physical activity and cardiovascular fitness with risk factors were found as in studies on less active populations.

The studies described in this dissertation can be classified as cross-sectional studies. A cross-sectional study is defined as one where the measurements of cause and outcome refer to the same or approximately the same point in time, and there is no inbuilt directionality. A common reason for this lack of directionality is that both postulated causal factor

and the postulated outcome have been present for some time before the study, and it is not certain which came first (the chicken-or-egg problem). A study of relationship between obesity and physical inactivity, measured at the same time, is an example. Also, in some instances it may not be conceptually clear that one of the variables can be considered a cause of the other. Still, the results of our studies point in the same direction as general accepted findings in the literature.

Obviously due to the selection criteria, the study population can be regarded as a positive selection according to health parameters compared to the general population of the same age. The findings in our studies confirmed again the important point, as all large-scale investigations, that adequate physical activity is essential to cardiovascular health. Studies of physical fitness also point in the same direction as the physical activity findings. The relationships between the level of habitual physical activity, fitness and health are very complex. Other factors are associated with individual differences in health status. Likewise, the levels of physical fitness are not determined entirely by an individual's level of habitual physical activity. Other lifestyle components, environmental conditions, personal attributes, and genetic characteristics also affect the major factors: physical activity, fitness, and health. In view of this complex model one can imagine for example, that there are individuals who are gifted with a relatively high level of cardiovascular fitness by genetic endowment, whereas others have to train regularly to achieve the same effect of protection against cardiovascular diseases. There is an indication, in view of the great variation in characteristics and responses to exercise, that even little changes in habitual physical activity may contribute to changes in risk profiles.

Furthermore, an important aspect of cost-effectiveness studies on physical activity is the fact that a considerably large proportion of the general population is physically inactive. This means that the impact of physical inactivity on cardiovascular disease is considerably larger than that of most of the other known risk factors. Therefore, much more effect can be expected from health promotion programmes on active lifestyles. These health promotion activities should be focused especially on those who are not yet habitually active. It is also important that these potential entrants into a programme do not dislike exercise because a certain level of enjoyment is desired for the continuation of habitual physical activities.

A crucial role in the promotion of habitual physical activity has to be reserved for the general practitioner. He is best informed on the present

health status of his patients and he can select and motivate potential candidates for exercise programmes. One may ask which organization can provide for the implementation of such programmes. It seems obvious that Centers for Public Health have potential knowledge to take care of large-scale health problems. They can initiate activities varying from health education to a preventive sports-medical screening. We have to be careful not to confound physical activity and sport. As was concluded in our study, those who volunteered in a sports-medical screening are not representative for a physically active population. There is need for a medical screening with those who have intentions to start or continue with strenuous exercise. In the first place many people are afraid of negative effects of physical activity or are not aware of positive effects. Individual advices on basis of results of a screening will reduce this fear and ignorance. Although not many abnormalities with potential risks during exercise are to be expected in an asymptomatic population, the few subjects at risk may be identified. Furthermore, the results of the examination give an impression of individual risk profiles and an opportunity to compile a suitable individual exercise programme for the majority. Another advantage of documenting information on the individual health status is the possibility to evaluate effects of these programmes, which can be important for the motivation of individuals to continue with habitual physical activities. The costs of this approach need not be very high, moreover, the benefits of habitual physical activity (e.g., reduced demands for medical care, increased industrial productivity, and decreased absenteeism) can outweigh these costs.

Samenvatting

In dit proefschrift worden resultaten besproken van testen die zijn uitgevoerd in het kader van een grootschalig onderzoek naar de effectiviteit van de medische sportkeuring. Het onderzoek vond plaats op verzoek van het ministerie van WVC en had tot doel gegevens te verzamelen voor beleidsmakers op het gebied van de preventieve (sport)-geneeskunde. Tussen 1984 en 1988 werden bijna 4000 sporters van 40 jaar en ouder uitgebreid onderzocht nadat ze zich vrijwillig hadden aangemeld voor een sportkeuring. Het doel van de studie was vast te stellen hoeveel (sport-)relevante afwijkingen gevonden konden worden en tot welke adviezen deze onderzoeken aanleiding gaven. Alle onderzochte personen werden na een en na twee jaren gevraagd wat er in de tussenliggende periode gebeurd was naar aanleiding van de keuring en of ze zich de gegeven adviezen nog konden herinneren. De resultaten van deze studie zijn deels reeds gepubliceerd in een rapport en in diverse artikelen door de verschillende onderzoekers. Andere publicaties van deze studie zijn nog in voorbereiding. In deze dissertatie zijn de artikelen gebundeld van één van de onderzoekers. In zijn geheel kunnen deze artikelen van belang zijn voor de interpretatie van testresultaten van de onderzoekspopulatie en kunnen tevens dienen voor de vergelijking met gegevens van andere onderzoeken. Omdat de keuringen op drie verschillende lokaties plaatsvonden, drie GGD-en in Zuid-Limburg, moest rekening gehouden worden met gevolgen voor de betrouwbaarheid van de testresultaten. Verder bleek dat vergelijkingen met referentiewaarden, meestal waarden van de doorsnee bevolking, in het voordeel van de onderzoekspopulatie uitvielen omdat we te maken hadden met een relatief gezonde en actieve selectie. Het was eveneens interessant om voor deze actieve 40-plussers te onderzoeken hoe de relaties tussen fitheid, lichamelijke activiteit en enkele risicofactoren voor hart- en vaatziekten waren.

Hoofdstuk 1 kan gezien worden als een algemene introductie op de overige hoofdstukken. Het geeft een beschrijving en resultaten van die variabelen van de studie die van belang zijn voor de interpretatie van resultaten in de volgende hoofdstukken. De totale onderzochte groep van 3918 vrijwilligers bestond uit 2661 mannen en 1257 vrouwen. De mannelijke sporters waren gemiddeld 4,3 uren per week bezig met sport, terwijl de vrouwen gemiddeld 3,5 uren per week sportten. Het protocol van het sportmedisch onderzoek was zeer uitgebreid en behelsde een vragenlijst over klachten en ziektegeschiedenis, inclusief blessures,

alsmede huidige en vroegere (sport-)activiteiten, een uitgebreid lichame-
lijk onderzoek, laboratoriumtesten, inclusief een maximale inspannings-
test op een fietsergometer en een evaluatie van de testresultaten met
zowel sportgerichte als algemene gezondheidsadviezen. Vergeleken met
een representatieve steekproef van mannen en vrouwen uit dezelfde
leeftijdsgroep in Zuid-Limburg, was deze onderzoeksgroep jonger en
actiever dan het sportende deel van de steekproef. Verder bestond de
groep uit relatief meer mannen en beoefenden zowel mannelijke als
vrouwelijke sporters vaker sporten waarbij een hoger inspanningsniveau
wordt bereikt. Dit zou mogelijk een verklaring kunnen zijn voor de
hogere blessure-incidentiecijfers vergeleken met sporters uit de eerder
vermelde representatieve steekproef.

In hoofdstuk 2 en hoofdstuk 3 worden resultaten besproken van een
studie naar de betrouwbaarheid van metingen die door verschillende
onderzoekers (inter-waarnemer) op verschillende lokaties met eigen
apparatuur zijn uitgevoerd. Voor dit doel zijn vijf groepen van 9-10
vrijwilligers op de drie lokaties onderzocht. In totaal zijn in een periode
van 2,5 jaren, 43 mannen en 5 vrouwen binnen drie weken driemaal
gekeurd. **Hoofdstuk 2** handelt over de variabiliteit en betrouwbaarheid
van flexibiliteitsmetingen, uitgevoerd met een goniometer. Deze methode
wordt op grote schaal toegepast om een indruk te krijgen van bewegings-
uitslagen van gewrichten. Het geheel van deze flexibiliteitsmetingen
wordt gebruikt als maat voor de lenigheid. In de kliniek wordt vaak met
behulp van goniometrie het resultaat van een behandeling vastgesteld.
Voor de flexibiliteitsmetingen werd nog een aparte studie uitgevoerd bij
acht universiteitsmedewerkers om tevens een indruk te krijgen van de
variabiliteit en betrouwbaarheid van dit soort metingen binnen een
onderzoeker (intra-waarnemer). De resultaten laten zien dat het moeilijk
is om bij een individu een toename of een afname van gewrichtsuitlagen
van minder dan 5 tot 10 graden aan te tonen voor de meeste gewrichten,
gemeten door dezelfde onderzoeker. De inter-waarnemer variabiliteit
bleek niet consistent te zijn over de gehele onderzoeksperiode, met als
gevolg dat verschillen tussen waarnemers niet gecorrigeerd kunnen
worden op basis van metingen op een bepaald tijdstip. Verder bleek dat
de betrouwbaarheid van de meeste flexibiliteitsmetingen verbeterde wan-
neer de ervaring van de onderzoekers toenam.

In **hoofdstuk 3** worden de resultaten van de betrouwbaarheidsstudie
besproken van drie longfunctietesten (vitale capaciteit, één-seconde
waarde en de "peakflow" waarde), een aantal metingen tijdens de
maximale inspanningstest op een fietsergometer (maximale vermogen,

maximale systolische bloeddruk en maximale hartfrequentie), de bloeddruk in rust en het vetpercentage, geschat aan de hand van de som van 4 huidplooien. De resultaten laten zien dat de variatie tussen verschillende waarnemers relatief gering is ten opzichte van de variatie tussen individuen zodat de gegevens, die op verschillende lokaties verzameld zijn, gezamenlijk gebruikt kunnen worden in een transversale studie. De biologische variatie lijkt de belangrijkste bron voor de totale variatie tussen laboratoria bij maximale inspanningstesten.

Het doel van de studie beschreven in **hoofdstuk 4** was te onderzoeken in hoeverre gezonde en actieve 40-plussers betere longfunctiewaarden hebben dan voorspelde normaalwaarden. Bij jongere atleten werd reeds vastgesteld dat zij aanzienlijk betere longfuncties hadden dan niet actieve leeftijdsgenoten. Het zou dus kunnen gebeuren dat een sporter een behoorlijk deel van zijn longfunctie verloren heeft gedurende een bepaalde periode zonder dat dit uit de vergelijking met referentiewaarden zou blijken, hij heeft immers nog normale waarden. De gevonden verschillen van de vitale capaciteit (VC), de geforceerde vitale capaciteit (FVC), de één-seconde waarde (FEV1.0) en de peakflow tussen waarden van de onderzoeksgroep en referentiewaarden waren aanzienlijk, zodat nieuwe regressievergelijkingen werden berekend, die gebruikt kunnen worden voor de voorspelling van longfunctiewaarden van gezonde en actieve mensen in deze leeftijdsgroep. Bij deze berekeningen werd rekening gehouden met rookgewoonte. De FEV1.0/VC ratio leverde geen verschillen op met referentiewaarden. Verder bleek dat sporters, die duursporten als joggen, wielrennen, roeien, schaatsen of langlaufen beoefenden, hogere longvolumes (VC, FVC en FEV1.0) hadden dan de overige sporters.

De maximale inspanningstest op een fietsergometer is belangrijk om een indruk te krijgen van de cardiopulmonaire reactie op inspanning. Het is verder een betrouwbare methode om het aerobe uithoudingsvermogen (W_{max}) te bepalen, een maat voor de lichamelijke fitheid. De onderzochte sporters werden aangemoedigd om tijdens de test met een opklimmende belasting tot uitputting door te fietsen. Om objectief te beoordelen of iemand zich maximaal had ingespannen, werd in de derde minuut na inspanning een veneus bloedmonster afgenomen voor de bepaling van het "peak" plasmalactaatgehalte (PPLa). De maximale hartfrequentie en PPLa zijn goede alternatieven voor de beoordeling van een maximale belasting, wanneer geen apparatuur voor de bepaling van de zuurstofopname beschikbaar is. In **hoofdstuk 5** wordt beschreven dat er een verschil van 2,5 mmol/l in PPLa werd gevonden tussen mannen en

vrouwen, onafhankelijk van leeftijd en maximale hartfrequentie. Dit gegeven is gebruikt bij het vaststellen van objectieve criteria voor een maximale inspanning van de onderzoeksgroep.

Zowel in klinische situaties als in de sportmedische praktijk wordt in toenemende mate gebruik gemaakt van de inspanningstest. De fietsergometer is hierbij een valide instrument om personen bij verschillende intensiteiten te belasten. In hoofdstuk 6 worden referentiewaarden voor parameters van de maximale inspanningstest zoals W_{max} , PPLa en maximale hartfrequentie en systolische bloeddruk gepresenteerd voor mannelijke en vrouwelijke sporters van middelbare leeftijd. Verder is de relatie bestudeerd tussen W_{max} en onafhankelijke variabelen als leeftijd, antropometrische parameters, parameters van lichamelijke activiteit, rookgewoonte, vitale capaciteit en hartfrequentie. Met behulp van multiële regressie analyse werd aangetoond dat ongeveer de helft van de variatie in W_{max} verklaard kan worden door de onafhankelijke variabelen. De resultaten geven een indruk over de relatieve bijdrage van verschillende factoren aan het maximale prestatievermogen waarbij opvalt dat lichamelijke activiteit naast de sportbeoefening, zoals het gebruik van de fiets als transportmiddel, een significante extra bijdrage levert aan het prestatievermogen. Bij de vrouwelijke sporters is deze bijdrage meer uitgesproken.

Diverse onderzoeken hebben reeds de relatie beschreven tussen lichamelijke activiteit en risicofactoren voor hart- en vaatziekten of de relatie tussen lichamelijke fitheid en deze risicofactoren. Zelden zijn deze beide relaties binnen dezelfde onderzoeksgroep bestudeerd. In hoofdstuk 7 is onderzocht hoe de relatie is tussen lichamelijke fitheid, in dit geval W_{max} , lichamelijke activiteiten en enkele cardiovasculaire risicofactoren als bloeddruk, lichaamssamenstelling en rookgewoonte, in een oudere, gezonde en actieve populatie. De associatie tussen lichamelijke activiteiten en risicofactoren was minder sterk dan die tussen fitheid en risicofactoren, bij zowel mannen als vrouwen. Uit correlaties bleek dat sporters met een betere fitheid gunstigere risicoprofielen hadden. Hogere correlaties tussen fitheid en risicofactoren werden gevonden bij vrouwen vergeleken met mannen, behalve voor rookgewoonte. Wanneer de sporters werden ingedeeld naar type van sportbeoefening bleek dat zowel mannelijke als vrouwelijke joggers de meest gunstige risicoprofielen hadden, vergeleken met beoefenaren van één van vier andere, meest populaire sporten. Bij de mannen waren deze sporten wielrennen, tennis, voetbal en volleybal en bij de vrouwen tennis, gymnastiek, zwemmen en wandelen. De sterkte van de associaties tussen lichamelijke fitheid en

activiteiten enerzijds en risicofactoren anderzijds in onze relatief actieve en gezonde onderzoeksgroep zijn vergelijkbaar met de gevonden associaties bij minder actieve populaties.

In **hoofdstuk 8** worden de belangrijkste resultaten nog eens samengevat waarna een algemene discussie volgt over de invloed van lichamelijke activiteiten op de gezondheid. Omdat deze invloed zowel op individueel als groepsniveau belangrijke gevolgen heeft, lijkt een samenwerking tussen diegenen die kennis hebben van individuele gezondheidsproblemen, d.w.z. huisartsen en organisaties op het gebied van de collectieve preventie (o.a. GGD-en) het meest voor de hand te liggen om gunstige effecten te bewerkstelligen. Bij deze samenwerking kan adequaat omgegaan worden met aspecten als het nut van preventief screenen op risicofactoren op individuele basis, het motiveren van potentiële deelnemers aan bewegingsprogramma's, evaluatie van bewegingsprogramma's, alsmede preventie van sportblessures.

Nawoord

Terugkijkend op de jaren die ik in dienst van het Instituut Sportgeneeskunde Limburg, gedetacheerd bij de Rijksuniversiteit Limburg, heb gewerkt, krijg ik naast een gevoel van weemoed tevens gevoelens van dankbaarheid tegenover al diegene, die voor een optimaal werkklimaat zorgden. Zij maakten het mogelijk om mijn wetenschappelijk activiteiten af te ronden met een promotie. Zelfs in het laatste jaar, toen ik reeds bij de GGD Westelijke Mijnstreek werkzaam was in een nieuwe functie, vormden zij de stimulans om door te zetten.

Een woord van dank is dan ook hier op zijn plaats voor al diegenen die hieraan meewerkten, waarbij ik enkelen in het bijzonder zou willen noemen, beseffende dat dit altijd risico's inhoudt anderen tekort te doen. Allereerst wil ik mijn mede-onderzoekers Sjef Vrencken en Sjef Wijnen bedanken, niet alleen voor de mede uitvoering van het 40+ onderzoek, maar meer nog voor de prettige wijze waarop we al die jaren hebben samengewerkt. Dat we al die jaren een hecht team gevormd hebben moge blijken uit het feit dat ik nog regelmatig met Sjef wordt aangesproken. Ik hoop dat onze vriendschap voor altijd gestalte heeft gekregen.

Vervolgens wil ik de grote inbreng van beide co-promotores, Marleen van Baak en Frans Verstappen, noemen. Frans heeft er niet alleen voor gezorgd, dat de financiële middelen voor het onderzoek beschikbaar kwamen, maar is tevens de stuwende kracht bij de uitvoering van het 40+ onderzoek en tal van andere onderzoeken geweest. Ik denk met veel plezier terug aan onze samenwerking. Marleen wil ik bedanken voor al haar wijze raad bij mijn wankel schreden op het wetenschappelijke terrein. Haar bijdrage aan dit proefschrift was van essentieel belang. Zij wist altijd in enkele woorden mij op het goede spoor te zetten. Ik ben blij dat haar eindeloos geduld nu beloond wordt.

Mijn promotor, Wim Saris, ben ik zeer erkentelijk voor zijn waardevolle commentaar in de afrondingsfase van het proefschrift. Hij wist, ondanks zijn drukke agenda, ervoor te zorgen dat het strakke tijdsschema, dat we het laatste jaar hadden opgesteld, in de praktijk ook gerealiseerd kon worden. Prof. dr. Tj. de Boorder wil ik danken voor zijn bijdrage aan het onderzoek. In de beginfase was hij als projectleider betrokken bij de opzet en uitvoering van het onderzoek. Zijn "drie musketiers" konden verder altijd rekenen op de gastvrijheid van de gehele vakgroep Arbeidsgeneeskunde, Milieugezondheidkunde en Toxicologie.

Tijdens de praktische uitvoering van het 40+ onderzoek hebben tal van medewerkers van de drie GGD-en in Zuid-Limburg meegewerkt aan het

verzamelen van de onderzoeksgegevens. Voor mij persoonlijk was de assistentie van Marianne Hermans en Paula Claessens van de GGD in Geleen van grote waarde. Bij de universiteit droegen met name Peter Geurten en Gerrit van Kranenburg hun steentje bij met de lactaatbepalingen, terwijl Jacques Lambregts altijd klaar stond wanneer de apparatuur haperde.

Jolanda Nelissen-Braeken maakte met behulp van de computer de onderzoeksgegevens toegankelijk. Zij zorgde er tevens voor dat het kille onderzoeksklimaat van het instituut met een vleugje romantiek wat aangenamer werd. Goede herinneringen heb ik ook overgehouden aan de periode bij de vakgroep Bewegingswetenschappen. Ik hoop in de toekomst nog regelmatig met haar medewerkers te kunnen samenwerken. Speciale dank gaat uit naar Marion Wishaupt-de Vreede die door haar spontane hulp ervoor heeft gezorgd dat het manuscript een mooi boekje werd.

Ik ben me er terdege van bewust dat alles nooit mogelijk zou zijn geweest zonder de geweldige steun van mijn gezin. Martine, jouw bijdrage aan dit boekje zit niet alleen in het ontwerp van de omslag en de vele uren tekstverwerking, maar voornamelijk in de liefdevolle steun, die voor het ideale werkklimaat zorgde tijdens al die avonden en weekenden.

Curriculum vitae

1957	Geboren op 21 februari te Maastricht
1976	Atheneum B, Henric van Veldekecollege te Maastricht
1982	Doctoraalexamen, Rijksuniversiteit Limburg te Maastricht
1983	Artsexamen, Rijksuniversiteit Limburg te Maastricht
1983-1990	Wetenschappelijk medewerker Instituut Sportge- neeskunde Limburg gedetacheerd bij de vakgroep Arbeidsgeneeskunde, Milieugezondheidkunde en Toxicologie van de Rijksuniversiteit Limburg te Maastricht
1990-heden	Arts afdeling Algemene Gezondheidszorg van de GGD Westelijke Mijnstreek te Geleen